DNA replication with an unnaturally large replicon in *Saccharomyces cerevisiae*

Ana Barbara Alice Wallis

née *Anna Barbara Alice Clark*

St. Catherine's College, Oxford

Sir William Dunn School of Pathology, Medical Sciences Division

A thesis submitted to

the University of Oxford for the degree of

DPhil in Interdisciplinary Bioscience (Doctoral Training Programme),

Trinity Term 2019

Word count: 48611



UNIVERSITY OF OXFORD

Abstract

Eukaryotic cells co-regulate DNA replication and the cell cycle, ensuring a single round of replication occurs during each cell cycle. Until a few years ago, origin licensing and subsequent activation was considered to be a stepwise activation process, dependent on cell cycle kinases. However, it is now known that some phosphatases also play a key role in regulating DNA replication. A potential role during DNA replication initiation for the phosphatase subunit Rts1 was examined. Temperature-sensitive replication initiation mutants were examined to determine if synthetic rescue resulted when *RTS1* was deleted (indicating antagonistic functions). It was shown that Rts1 did not antagonise origin licensing or activation. Conversely, *RTS1* was found to have negative genetic interactions with *CDC9*, *DBP11* and *CDC45*.

Re-initiation of replication origins (re-replication) has been associated with tumorigenesis; however, the opposite phenomenon (under-replication) has been less well studied. Potential outcomes include single-stranded DNA protection, alternative replication mechanisms, chromosome breaks, ‘fragile sites’ and/or cell death. This project aimed to create a system containing an unnaturally large replicon of over 100 kb, and characterise the resulting effects on cell growth and DNA replication.

DNA replication was delayed in the absence of adjacent active origins. However, mean generation time was not always extended. A next-generation sequencing approach, utilised to examine replication fork direction, indicated that replication forks could initiate close to telomeres despite origin inactivations, termed telomeric initiation.

Acknowledgments

First, I want to thank my primary supervisor, Conrad Nieduszynski. Conrad has been a huge source of support for my PhD, not only in terms of his scientific guidance but also in terms of all the motivational and inspirational support needed.

Additionally, I have been lucky enough to have the support of my secondary supervisor Chris Norbury this year, whose scientific advice and guidance whilst planning and writing up my thesis has been amazing – thank you very much.

A huge thank you is due to all the members of the Nieduszynski lab for providing absolutely everything I could have asked for – help, guidance, role models, sympathetic listening and most importantly friendship.

I absolutely cannot give enough credit and thanks to my husband, Charlie Wallis, for his huge support and patience throughout my course, including (but not limited to) cutting short holidays to visit yeast cells, late-night lab visits to keep me company, and mid-day phone calls for moral support.

Thank you to all of my wonderful family and friends for their support in helping Charlie and I during our whole lives, and especially the last couple of years. I cannot imagine getting here without all of your kindness, generosity and encouragement – you are absolute stars!

Finally, I would like to thank our collaborators and other laboratories for their help during my DPhil. Thank you to Tomoyuki Tanaka, David Shore, Anne Donaldson, Carol Newlon and Toshio Tsyukiyama for sending me yeast strains. I am also grateful to Tony Carr and his lab for hosting Phoebe and me to learn Polymerase-usage sequencing (Pu-seq).

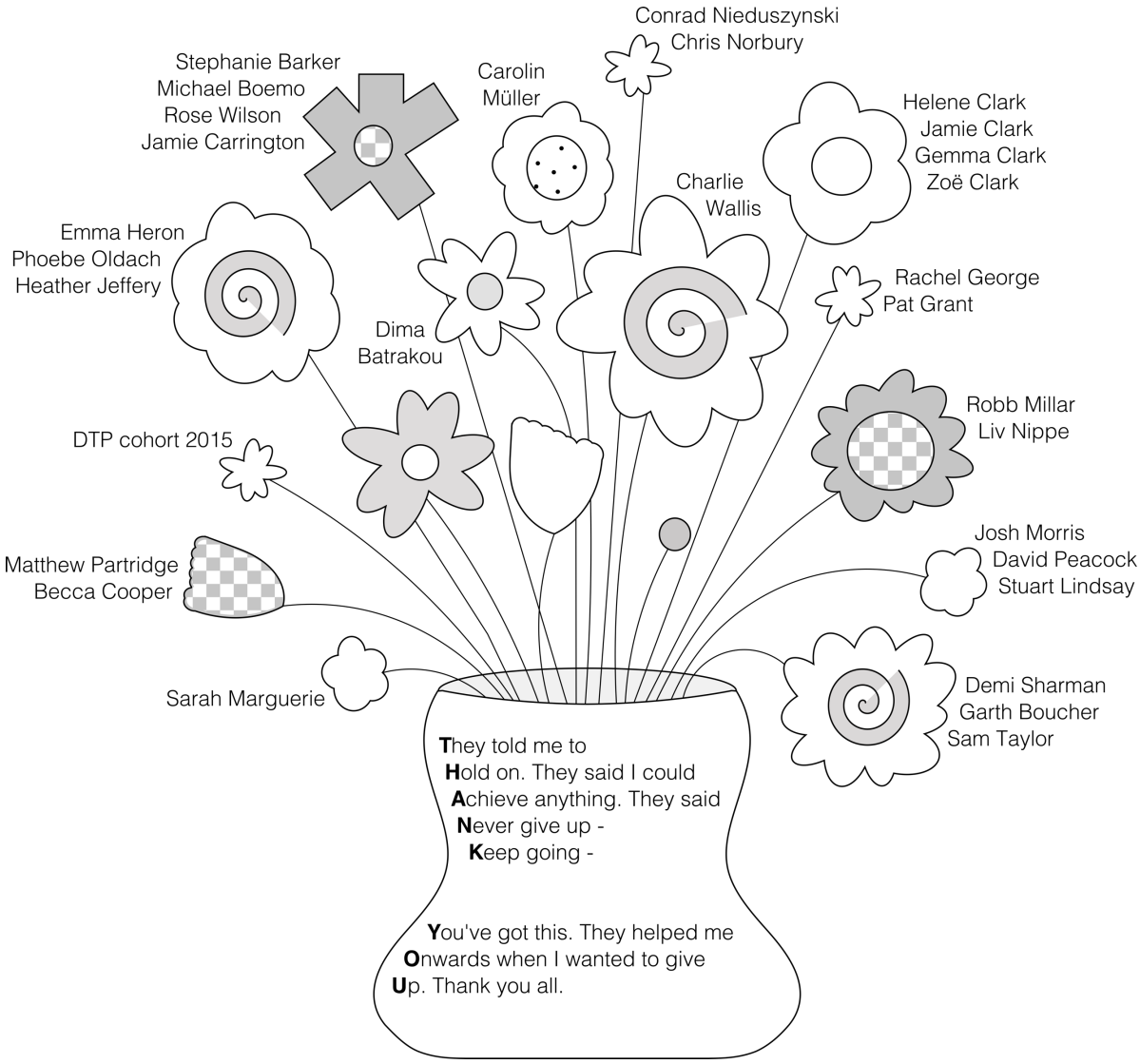


Figure 0.1 Thank you all. *A huge number of people have helped me through my DPhil, and there aren't enough words to express my gratitude to all of them. This image represents the names of all those I would like to particularly thank for all of their support, encouragement and guidance, without which I could not have completed the last four years.*

Abbreviations

AAA^{+}	ATPase associated with diverse cellular activities
ACS	ARS consensus sequence
ARS	Autonomously replicating sequence
APC	Anaphase promoting complex
BAH	Bromo-adjacent homology domain
BIR	Break-induced replication
bp	Base pair
CDK	Cyclin-dependant kinase
CEN	Centromere
ChIP	Chromatin immunopreciopitation
Chr	Chromosome
CKI	Cyclin dependent kinase inhibitor
CMG	Cdc45-Mcm2-7-GINS complex
$CRL2^{Lrr1}$	Cullin RING Ligase 2 associated with Lrr1
cSDR	Constitutive stable DNA replication
CTR	Constant timing region
DDK	Dbf4-dependant kinase
ddPCR	Droplet digital PCR
Dia2	Digs into agar 2
DSB	Double strand break
dsDNA	Double stranded DNA
DUE	DNA unwinding element
eACS	Extended ACS
Emi1	Early mitotic inhibitor 1
EDTA	Ethylene diamine tetraacetic acid

FACS	Fluorescence-activated cell sorting
FACT	Facilitates chromatin transcription
Fob1	Fork blocking protein 1
GINS	Sld5, Psf1, Psf2 and Psf3
HAT	Histone acetyltransferase
HDAC	Histone deacetylase complex
HoRReR	Homologous-recombination-restarted-replication
iSDR	Induced stable DNA replication
kb	Kilo base pair
LB	Luria-Bertani (broth/media)
Lrr1	Leucine Rich Repeats 1
Mb	Mega base pair
MCM	Mini-chromosome maintenance protein complex (Mcm2-7)
NEB	New England Biolabs
OD_{600}	Optical density at 600 nm
OGRE	Origin G-rich repeat element
ORC	Origin recognition complex (Orc 1-6)
ORB	Origin recognition box
OriDB	DNA replication origin database (www.oridb.org/)
Pol	Polymerase
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PP1	Protein phosphatase 1
PP2A	Protein phosphatase 2A
PPP	Phosphoprotein phosphatases
pre-LC	Preloading complex
pre-RC	Pre-replication complex

Pu-seq	Polymerase usage sequencing
qPCR	Quantitative PCR
rDNA	Ribosomal DNA
RFB	Replication fork barrier
RFC	Replication factor C
Rif1	Rap1 interacting factor
RPA	Replication protein A
Rrm3	rDNA recombination mutation protein 3
SCF^{Dia2}	Skp, Cullin, F-box containing complex associated with Dia2
SDS	Sodium dodecyl sulphate
SGD	Saccharomyces genome database
SGDP	Saccharomyces genome deletion project
SGRP	Saccharomyces genome resequencing project
ssDNA	Single stranded DNA
TAD	Topologically associated domain
TE	Tris EDTA
TERRA	Telomeric repeat-containing RNA
TIR	Transcription-induced replication
Top	Topoisomerase
tRNA	Transfer RNA
TTR	Timing transition region
UCSC	University of California, Santa Cruz (genome browser)
YPAD	Yeast-Peptone-Adenine-Dextrose

Table of contents

1 Introduction 1

1.1 The replicon model of DNA replication 1

1.2 DNA replication origins and initiators 1

1.2.1 Bacteria *2*

1.2.2 Archaea *4*

1.2.3 Yeast *6*

1.2.4 Metazoans *10*

1.3 Yeast as a model organism for DNA replication 12

1.3.1 The yeast genome *14*

1.3.2 The mitotic cell cycle in yeast *18*

1.4 Canonical DNA replication in yeast 20

1.4.1 Origin licensing (G1 phase) *20*

1.4.2 Origin firing (S phase) *24*

1.4.3 Replication elongation *28*

1.4.4 Replication termination *33*

1.5 Control of the cell cycle in yeast 40

1.5.1 Kinases *40*

1.5.2 Phosphatases *43*

1.6 Replication origin characteristics 45

1.6.1 Origin timing *46*

1.6.2 Origin distribution and efficiency *51*

1.7 Implications of de-regulated DNA replication 53

1.7.1 Over-replication *53*

1.7.2 Under-replication *55*

1.8 Alternative methods of DNA replication 56

1.8.1 Bacteria *57*

1.8.2 Archaea *59*

1.8.3 Budding yeast *61*

1.8.4 Other eukaryotes *68*

1.9 Objectives of this study 69

2 Methods 71

2.1 Chemicals and enzymes 71

2.2 Growth media 71

2.3 Yeast strains, plasmids and oligonucleotides 75

2.4 Molecular genetic methods 91

2.4.1 Chemical transformation of E. coli *cells* *91*

2.4.2 Plasmid extraction from E. coli *cells* *91*

2.4.3 DNA quantification *92*

2.4.4 Polymerase chain reaction (PCR) *93*

2.4.5 Restriction enzyme digestion of DNA *94*

2.4.6 DNA ligation *94*

2.4.7 Agarose gel electrophoresis *95*

2.4.8 DNA extraction from agarose gels *96*

2.4.9 Sanger DNA sequencing *96*

2.4.10 Ethanol precipitation of DNA *96*

2.4.11 Droplet Digital PCR (ddPCR) *97*

2.4.12 Preparation of polymerase-usage sequencing (Pu-seq) DNA *98*

2.4.13 Making DNA NextSeq libraries *100*

2.4.14 NextSeq DNA Sequencing *102*

2.5 Yeast methods 103

2.5.1 Growth conditions *103*

2.5.2 Calculating mean generation time *104*

2.5.3 Dilution series assay *104*

2.5.4 High efficiency yeast transformation *105*

2.5.5 ARS assay in yeast *106*

2.5.6 Colony PCR of yeast cells *106*

2.5.7 Mating type assays for yeast *106*

2.5.8 Mating yeast strains *107*

2.5.9 Harvesting and fixing yeast cells from liquid cultures *109*

2.5.10 Collection of stationary phase yeast cells *109*

2.5.11 Arresting yeast cells in alpha-factor *110*

2.5.12 Synchronous release of yeast cells for synchronous culture experiments *110*

2.5.13 'Smash and Grab' DNA extraction of yeast cells *111*

2.5.14 Measuring DNA content by flow cytometry *112*

2.5.15 Fluorescence activated cell sorting (FACS) of yeast *113*

2.5.16 DNA extraction of cells sorted using FACS *114*

2.5.17 Polymerase Usage sequencing (Pu-seq) *114*

2.6 Bioinformatics 116

2.6.1 Specialised computer programmes *116*

2.6.2 Web tools *117*

2.6.3 Unix-based tools *118*

3 Investigating genetic interactions between *RTS1* and replication initiation factors 124

3.1 Introduction 124

3.2 Generating *Δrts1* ts-origin factor double mutant strains 127

3.3 *RTS1* deletion does not rescue *orc2-1* temperature-sensitivity 131

3.4 *CDC6* does not interact genetically with *RTS1* 133

3.5 *Δrts1* does not antagonise or promote DDK phosphorylation 133

3.6 Deletion of *RTS1* increases the temperature-sensitivity of replication firing factor mutants.135

3.7 *RTS1* deletion increases the temperature-sensitivity of a *cdc9* ligase mutant.....137

3.8 *TRP1* deletion has a slight impact on temperature-sensitivity.....137

3.9 Summary 140

4 Characterising cell growth and DNA replication of *Saccharomyces cerevisiae* with a large replicon 144

4.1 Introduction 144

4.2 Characterising a disomic chromosome III strain 147

4.2.1 Disomic strains with inactive origins on the 174 kb ChrIII fragment grow more slowly ..147

4.2.2 DNA replication of a disomic chromosome III strain 150

4.3 Generating a large replicon on chromosome VI.....153

*4.3.1 Generating a four-origin deletion strain (*Δ4ori*) 153*

4.3.2 Confirming the ACS site of ARS608.5 .....*155*

4.3.3 Creating Δ6ori *large replicon strains.....160*

4.4 A large replicon on ChrVI does not affect cell growth.....162

4.6 DNA replication is delayed at inactivated ACS sites 168

4.7 High-resolution examination of DNA replication in S phase 175

4.8 Summary 180

5 Examining replication timing and polymerase usage across a large replicon in *Saccharomyces cerevisiae* 186

5.1 Introduction 186

5.2 Replication timing of late-replicating DNA can be resolved using synchronous culture experiments 193

5.3 Examining late-replicating DNA across a large replicon 197

5.4 Investigating polymerase-usage normalisation methods 203

5.4.1 Methods for normalising ribonucleotide incorporation data 203

5.4.2 Pol2-dependent replication in Δ4ori *cells 204*

5.5 Polymerase-usage sequencing of *Δ5ori* and *Δ6ori* cells 207

5.6 'Inactive' origins *ars606* and *ars609* appear capable of inefficient initiation 213

5.7 Modelling modes of ChrVI right replication 218

5.8 Summary 222

6 Discussion and conclusions 229

6.1 Discussion 229

6.2 Conclusions 241

7 Bibliography 244

8 Appendices 311

8.1 Pu-seq analysis script specific to this study 311

8.2 Wallis and Nieduszynski, 2018 317

1 Introduction

1.1 The replicon model of DNA replication

Simultaneous with the structure of DNA being discovered, a model for its replication was proposed: base-pairing nucleotides which transmit the genetic code (Watson & Crick 1953). However, to replicate the double-stranded DNA helix, a single-stranded template must be exposed, a step addressed by the 1963 replicon model; this proposed that to initiate DNA replication, a *cis*-acting replicator sequence would be bound by a *trans*-acting initiator protein, which would unwind the DNA helix (Jacob *et al.* 1963). Although originally proposed in *Escherichia coli*, this model has now been verified across all three domains of life by the discovery of DNA replication origins, although the *cis*-acting feature in mammalian cells is not thought to be sequence based.

1.2 DNA replication origins and initiators

Cells, in order to divide, must replicate their DNA. The site of DNA replication initiation is termed a replication origin, and the features that determine this site vary between the domains of life. In bacteria, archaea and budding yeast DNA replication origins are all defined by AT-rich DNA near conserved analogous sequences. In most eukaryotes, however, origins are not sequence specificed. These replicators recruit initiator proteins which lead to DNA replication initiation. However, although the principles of DNA replication are conserved between the domains, there are distinct differences between the three systems (reviewed in O'Donnell, Langston and Stillman, 2013) (Figure 1.1).

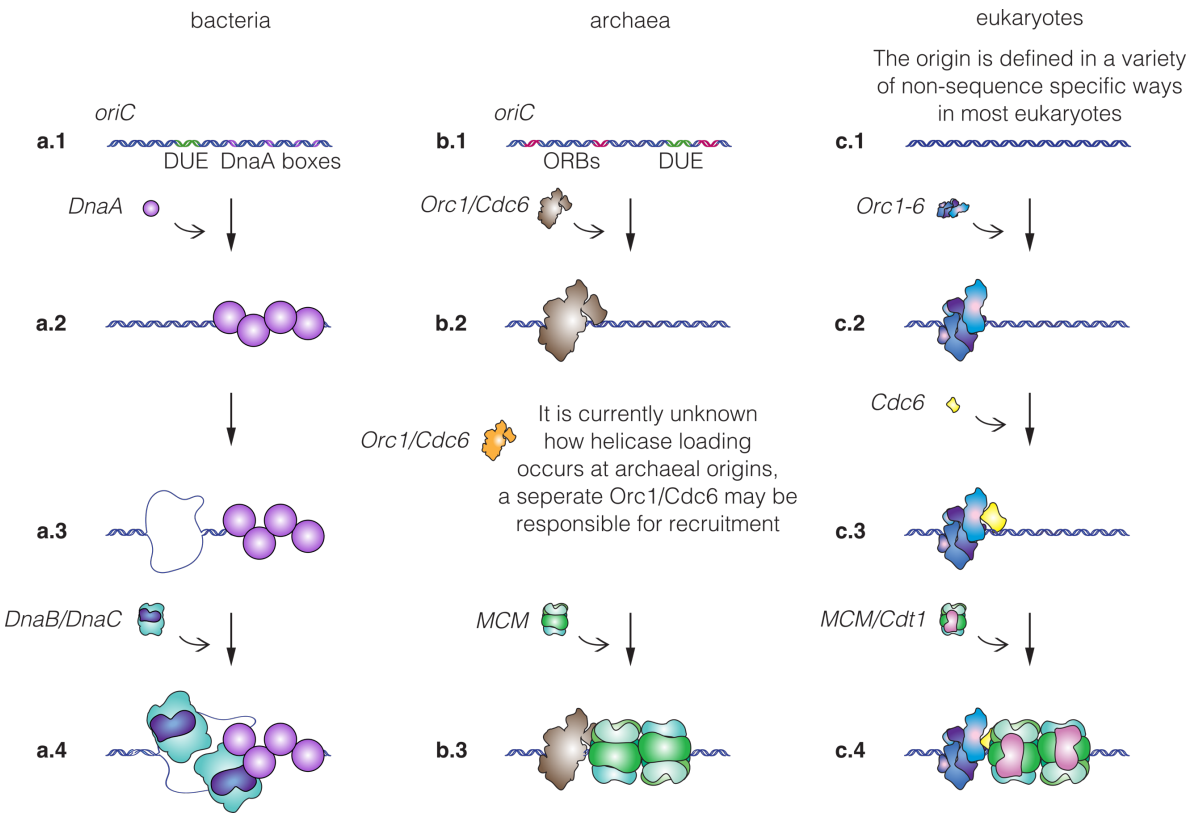


Figure 1.1 Schematic representation of origin sites and initiator proteins across three domains. *Bacteria, archaea and eukaryotes each have origin sites bound by an initiator protein (DnaA, Orc1/Cdc6 or Orc1-6 respectively). In* S. cerevisiae*, origins have a defined sequence but this is not conserved across eukaryotes. The initiator protein enables helicase binding, either DnaB in bacteria or the mini-chromosome maintenance complex (MCM) in archaea and eukaryotes. It is not known exactly how the archaeal helicase is recruited by the initiator protein. Adapted from (Leonard & Mechali 2013).*

1.2.1 Bacteria

Bacteria typically have one replication origin, known as *oriC*, per circular chromosome (Mackiewicz *et al.* 2004; Worning *et al.* 2006). The sites of replication origins have been predicted in more than 1300 bacterial genomes and

collated in a database and web-based search engine (Gao & Zhang 2007, 2008). The position of the bacterial origin is often next to the gene encoding the initiation protein, DnaA (Gao & Zhang 2007; Mackiewicz *et al.* 2004). The length of *oriC* varies between bacteria, from 100 – 1000 bp but it contains several conserved sites.

The first of these is the DNA unwinding element (DUE) which is AT-rich and facilitates DNA strand unwinding (Kowalski & Eddy 1989). Secondly, *oriC* contains a series of DnaA binding sites (DnaA boxes), which are often clustered within the *oriC* site and vary in their affinity for DnaA (Mackiewicz *et al.* 2004; Zawilak-Pawlik *et al.* 2005). For example, in the *E. coli oriC*, there are five DnaA boxes, three of which have high affinity for DnaA (Nievera *et al.* 2006; Schaper & Messer 1995). The three high affinity sites adhere to a consensus 9 bp DnaA box motif 5' TTATC/ACACA 3' whereas the lower affinity DnaA boxes differ by 2 bp or more from this sequence (Rozgaja *et al.* 2011). Finally, *oriC* contains binding sites for DNA bending proteins which alter the structure of *oriC*, altering DnaA binding affinity (Brassinga *et al.* 2002; Deog Su Hwang *et al.* 1992; Polaczek *et al.* 1997; Quon *et al.* 1998; Slater *et al.* 1995); for example the protein FIS represses initiation of DNA replication, whilst the IHF protein stimulates it (Ryan *et al.* 2004). Although bending protein sites are not required for cell viability, their loss does impact on the timing of pre-replication complex (pre-RC) assembly (Ryan *et al.* 2004).

The pre-RC in bacteria is composed of DnaA and DnaB, the DNA helicase. Once all DnaA sites are occupied in *oriC*, the DnaA complex unwinds DNA at the DUE element and a DnaB-DnaC complex is loaded (Kaguni 2011; Marszalek & Kaguni 1994). Despite DnaC sometimes being referred to as the 'helicase loader',

it is DnaA which recruits the DnaB-DnaC complex (Bárcena *et al.* 2001; Davey *et al.* 2002). DnaC has a more passive role, inhibiting DnaB function and it must leave the pre-RC in order for DnaB to function (Bárcena *et al.* 2001; Davey *et al.* 2002).

1.2.2 Archaea

Although archaeal chromosomes, like bacterial chromosomes, are circular, the archaeal replication machinery is more similar to that used by eukaryotes, with recent advances showing a remarkable level of conservation between the two systems (Makarova & Koonin 2013). The first archaeal origin was discovered in a *Pyrococcus* species, which has one origin per chromosome (Myllykallio *et al.* 2000). All archaeal origins share an AT-rich DUE element, similar to that in bacteria, and conserved repeats (termed origin recognition boxes, or ORBs) which are bound by the initiator protein (Wu *et al.* 2014).

However, unlike bacterial origins (Aves *et al.* 2012; O'Donnell *et al.* 2013), there is a huge diversity of origins within the archaea (Raymann *et al.* 2014). The number of ORBs and their sequence and orientation vary between archaeal genera (reviewed in Wu *et al.*, 2014). ORB sequence specificity appears to be less important for origin recognition in archaea than in bacteria. The two structures of origin recognition proteins isolated on DNA show only 4-5 bases of DNA are contacted by the origin recognition protein (Dueber *et al.* 2007). The number of origins per chromosome also varies between both genera and species. Some species have three origins per chromosome (Hawkins *et al.* 2013b; Lundgren *et al.* 2004; Norais *et al.* 2007), while one species, *Pyrobaculum calidifontis*, has four origins per chromosome, the largest origin number per

chromosome in a prokaryote recorded to date (Pelve *et al.* 2013). However, the number of origins per chromosome does not correlate with archaeal genome size (Figure 1.2) (Ausiannikava & Allers 2017).

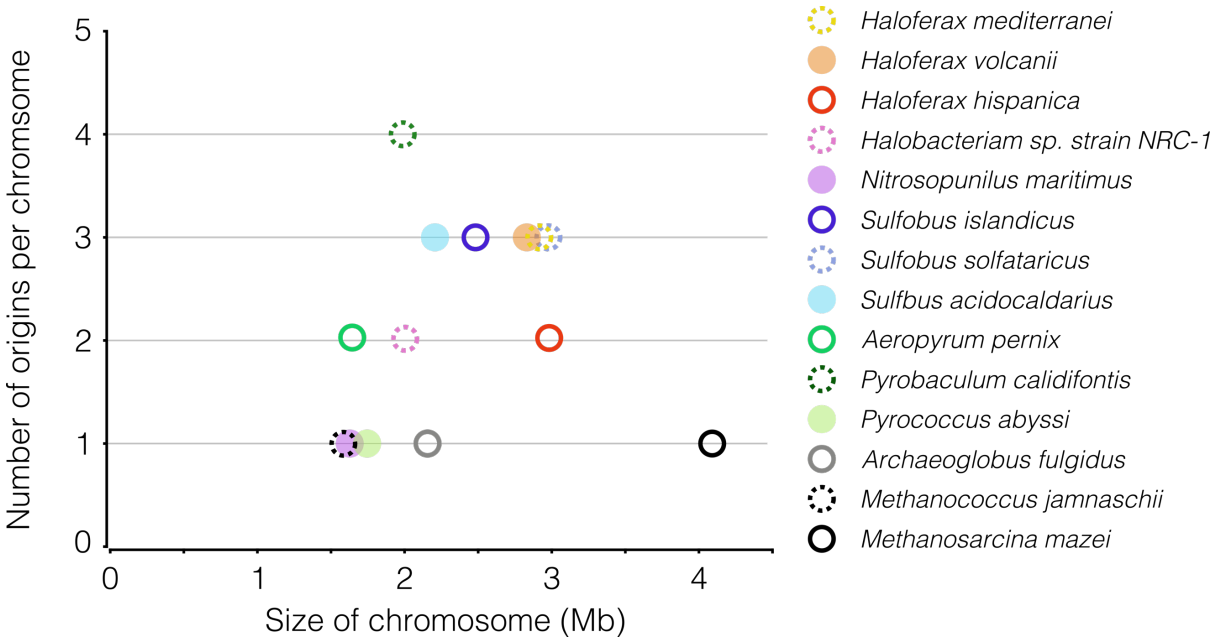


Figure 1.2 The number of origins per chromosome does not correlate with chromosome size in archaea. *Data taken from (Ausiannikava & Allers 2017).*

As in bacteria, each origin in archaea is associated with a gene encoding the origin recognition protein. The initiator protein in archaea is known as either Orc1 or Cdc6 depending on species, as the initiator protein shares homology with both the eukaryotic Orc1 and Cdc6 proteins (Ausiannikava & Allers 2017). The confusion over naming has been exacerbated by the diversity among Orc1/Cdc6 proteins in archaea (Makarova & Koonin 2013); one origin in *Sulfolobus islandicus* is bound by an alternative protein altogether (Samson *et al.* 2013). Orc1/Cdc6 proteins have been found to bind as monomers (Gaudier *et al.* 2007),

dimers (Grainge *et al.* 2006) and in some cases heterodimers between separate Orc1 and Cdc6 orthologues (Dueber *et al.* 2007).

How Orc1/Cdc6 proteins facilitate the process of DNA unwinding is not agreed upon, however unlike in bacteria unwinding does not appear to occur at the DUE element (Bell 2012). Additionally, since the number of Orc1/Cdc6 genes in a genome is often greater than the number of origins, archaeal Orc1/Cdc6 proteins may have roles other than in DNA replication initiation in archaea (Hartman *et al.* 2010; Hawkins *et al.* 2013b; Norais *et al.* 2007). Most archaea genomes contain one or two slow evolving (core) Orc1/Cdc6 genes, while other copies (shell copies) are evolving more rapidly (Raymann *et al.* 2014). Where two core copies exist, typically only one of these is linked to an origin and this gene is essential for replication initiation (Samson *et al.* 2013). However, it has been suggested that the second core copy in these instances may be able to bind origins and potentially act as a negative regulator of DNA replication (Robinson *et al.* 2004; Robinson & Bell 2007).

1.2.3 Yeast

1.2.3.1 *Saccharomyces cerevisiae*

In the budding yeast *Saccharomyces cerevisiae* DNA replication origins are termed autonomously replicating sequences (ARSs) due to their ability to support the replication of plasmids independently of other DNA sequence (Stinchcomb *et al.* 1979). The first yeast origin identified, *ARS1*, was identified as an 850 bp sequence, however, a consensus 11 bp core sequence required for replication, termed the ARS consensus sequence (ACS) was soon identified: 5'-[A/T]TTTAT[A/G]TTT[A/T]-3' (Palzkill *et al.* 1986). Each base (with the possible

exception of the T at position 6, which can be altered to a C) is required for origin function (Van Houten & Newlon 1990). This sequence was later expanded to an extended 17 bp ACS sequence (eACS), with additional nucleotides 3' to the T-rich strand; in *ARS309* two base pair changes are observed relative to the 11 bp consensus, however, this origin is surprisingly able to function, presumably because it adheres broadly to the eACS (Palzkill & Newlon 1988; Theis & Newlon 1997). The eACS sequence has been used to predict sites of *Saccharomyces cerevisiae* origins which are chromosomally active (Breier *et al.* 2004). Crucially, when mutations to the predicted ACS sequence are made in chromosomal origins, origin function is ablated (Deshpande & Newlon 1992; Rivier & Rine 1992)

The ACS is essential but not sufficient for replication (Bouton & Smith 1986; Brewer & Fangman 1987; Kearsey 1983). Three additional elements, referred to as B1, B2 and B3 (collectively the B domain) are also found at origins, 3' of the ACS, and the ACS with these three elements but no other flanking origin sequences are able to support replication (Figure 1.3) (Marahrens & Stillman 1992). A DUE sequence is also found 3' of the ACS, within the B domain (Umek & Kowalski 1988). It has been suggested that more flexibility in the sequence elements 3' of the core ACS is tolerated as their main function is to be easily unwound (Natale *et al.* 1992). Another school of thought was that the B domain contains elements for binding factors that regulate replication initiation. For example, the B3 element is a binding site for *ARS*-binding factor 1 (Abf1p) (Newlon & Theis 1993). These concepts are not mutually exclusive, and given the diversity of sequence observed in B elements relative to the ACS, it is now thought that the relative role of the B domain may vary between origins.

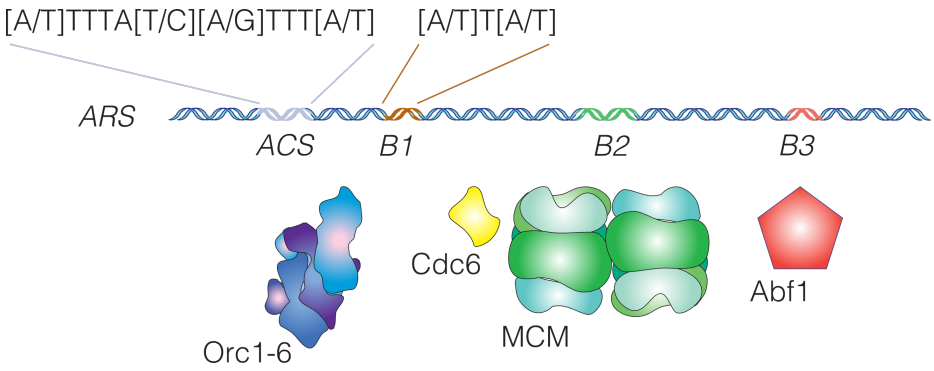


Figure 1.3 Schematic of* S. cerevisiae *origin conserved sequence elements. *There are multiple conserved sequence elements in* S. cerevisiae *origins, and different proteins bind to these elements. Orc1-6 binds the conserved 11 bp ACS and overlaps the B1 element. Cdc6 binds between the B1 and B2 elements. The MCM helicase is loaded onto the B2 element, and the B3 element is bound by Abf1. Adapted from (Leonard & Mechali 2013).*

The initiator protein in *S. cerevisiae* is the origin recognition complex (ORC) comprised of the Orc1-6 proteins (Bell & Stillman 1992). ORC binds over the ACS and B1 elements of the ARS, in an ATP-dependent manner, both *in vitro* and *in vivo* (Bell & Stillman 1992; Diffley & Cocker 1992). The binding of ORC limits replication of origins to once per cell cycle (Diffley *et al.* 1994). In eukaryotes, chromatin structure also plays a role in DNA replication initiation. ORC binding requires DNA to be nucleosome-free, helping to explain why there are more ACS matches than ORC binding sites in the *S. cerevisiae* genome (Berbenetz *et al.* 2010; Eaton *et al.* 2010). It has been shown that nucleosomes surrounding active origins are more dynamic than those elsewhere in the genome (Dion *et al.* 2007) and that origin function in a plasmid is impaired when local nucleosomes are moved closer to the ACS (Simpson 1990). How ORC recruits

the pre-RC and leads to DNA unwinding will be discussed in detail below (section 1.4).

1.2.3.2 Other yeast species

In contrast to *S. cerevisiae*, other yeast species (both budding and fission) do not have a specific conserved sequence required for replication, although several budding yeast species have ACS-like sequences (Liachko *et al.* 2010; Xu *et al.* 2012). For example, one budding yeast species, *Lachancea kluyveri* has been shown to contain a 9 bp sequence which is similar to the *S. cerevisiae* ACS, however it is not sufficient for ARS activity (Liachko *et al.* 2010). However, the greatest predictor for ARS activity in *L. kluyveri* was shown to be AT-richness (Liachko *et al.* 2011).

Origins in the fission yeast *Schizosaccharomyces pombe* are also AT-rich, with 69 – 75 % AT residues (Maundrell *et al.* 1988). Initially, origins were expected to be similar to ARS regions of *S. cerevisiae* in requiring specific sequence elements (Caddle & Calos 1994), and a putative 11 bp sequence for *S. pombe* origins was suggested as 5'-[A/T]PuTTTATTTA[A/T]-3' (Maundrell *et al.* 1988). However, it was also noted that mutating this sequence had little effect on ARS activity (Maundrell *et al.* 1988).

Consequently, replication origins in *S. cerevisiae* and *S. pombe* were found to be fundamentally different (Clyne & Kelly 1995). Unlike in *S. cerevisiae*, origins in *S. pombe* are large (> 1 kb) and do not have consensus elements required for initiation (Segurado *et al.* 2003). Instead, origins in *S. pombe* appear to comprise a series of non-essential AT-rich motifs which contribute cumulatively to ORC recruitment (Cotobal *et al.* 2010; Dai *et al.* 2005; Dubey *et al.* 1994;

Johnston & Barker 1987; Okuno *et al.* 1999). However, like *S. cerevisiae*, other sequence elements present at the origin may contribute to replication by recruiting enhancers or aiding with ORC recruitment (Okuno *et al.* 1999). Further, it has been shown that in *S. pombe* origins bind ORC as the initiator protein, and are kept nucleosome-free (Hayashi *et al.* 2007; Xu *et al.* 2012).

Finally, in the yeast species *Schizosaccharomyces japonicus* origins are GC-rich as opposed to AT-rich (Xu *et al.* 2012). However, the GC-rich origins are proposed to recruit ORC in a similar fashion to AT-rich regions, as it has been shown that poly(G) motifs also exclude nucleosomes (Fenouil *et al.* 2012; Tsankov *et al.* 2011). Although there are no precise sequence-defined origins in *S. pombe*, their ability to recruit ORC and initiate DNA replication is reproducible between cell cycles, giving a characteristic replication time for each ARS (Kim & Huberman 2001). Collectively, the diversity of origin structure amongst yeast species suggests that, in eukaryotes, the chromatin environment and DNA accessibility is more important than the precise DNA sequence at origin sites.

1.2.4 Metazoans

In multicellular eukaryotes, no conserved ORC-binding sequence has been determined (MacAlpine *et al.* 2010; Vashee *et al.* 2003). However ORC does act as the initiator protein, and this is conserved across metazoans; human ORC is able to restore replication activity to *Xenopus* egg extracts (Vashee *et al.* 2003). Metazoan ORC has also been shown to have roles in chromatin organisation aside from binding to replication origins (Hemerly *et al.* 2009; Pak *et al.* 1997; Prasanth *et al.* 2004). However, the origin-binding function of ORC is focussed on here in relation to metazoan replication origins.

Instead of a specific DNA sequence, the overall composition of DNA appears important for replication initiation, as in yeasts other than *S. cerevisiae*. Unlike most yeast, however, metazoan origins (including those in the model plant *Arabidopsis thaliana)* are GC-rich (Cadoret *et al.* 2008; Cayrou *et al.* 2011; Costas *et al.* 2011; Delgado *et al.* 1998; Sequeira-Mendes *et al.* 2009). Origins in hamster, human, mouse and *Drosophila* cells have been shown to be associated with CpG islands (stretches of DNA around one kb long with high CG content) (Cadoret *et al.* 2008; Cayrou *et al.* 2011; Delgado *et al.* 1998; Sequeira-Mendes *et al.* 2009). CpG islands may be methylated, however in *Drosophila* this is nearly absent suggesting methylation is not required for origin function (Cayrou *et al.* 2011). Origins often co-localise with active promoters; 85 % of replication sites in mouse embryonic stem cells were associated with transcriptional units (Sequeira-Mendes *et al.* 2009).

The secondary DNA structure may also play a role in some metazoan origins. Origin G-rich repeat elements (OGREs) are able to form G-quadruplexes, and are present in 67 % of *Drosophila* and 90 % of human replication origins (Cayrou *et al.* 2012). In both mice and *Drosophila*, replication initiation occurs approximately 200 bp away from the OGRE (Cayrou *et al.* 2012). Further, the torsion of the DNA is important, as while *Drosophila* ORC *in vitro* showed only a six-fold preference for origin DNA, it was 30 times enriched on negatively supercoiled DNA (Remus *et al.* 2004).

A collection of epigenetic marks are thought to specify some eukaryotic origins. Multicellular metazoan origins occur in nucleosome-free regions, where the open chromatin facilitates ORC binding (MacAlpine *et al.* 2010). Specific histone variants, such as H3.3 are also associated with origins (Jin *et al.* 2009;

MacAlpine *et al.* 2010; Stroud *et al.* 2012). The ability of ORC to associate with chromatin, as opposed to naked DNA, is highlighted by studies which show the bromo-adjacent homology domain (BAH) of Orc1 is required to activate replication origins in yeast and human cells (Müller *et al.* 2010; Noguchi *et al.* 2006). The BAH domain of Orc1 binds to methylated histone H4 (Kuo *et al.* 2012). Chromatin re-modellers have been shown to promote pre-RC recruitment and DNA replication (Miotto & Struhl 2010; Tardat *et al.* 2010). For example, histone acetylation in *Drosophila* cells promotes re-replication from origins during development (Aggarwal & Calvi 2004), and the histone methylation state of H3K4, H3K36 or H4K20 is important for activating or repressing origins, depending on the cell cycle state (Costas *et al.* 2011; Dorn & Cook 2011; Eaton *et al.* 2010).

Finally, larger scale interactions of chromatin domains might also contribute to DNA replication initiation (Chen *et al.* 2011). Loops of chromatin bound by cohesin have been suggested to be involved in organising replication timing domains (Cayrou *et al.* 2012; MacAlpine *et al.* 2010), and cohesin enrichment is observed at some origins (Guillou *et al.* 2010). Clusters of origins form replication factories in cells (Jackson & Pombo 1998), and these domains correlate with long-range chromatin interactions (Ryba *et al.* 2010). However, cohesin ablation does not impact on DNA replication timing suggesting cohesin-mediated genome organisation is not required to establish replication timing domains (Oldach & Nieduszynski 2019).

1.3 Yeast as a model organism for DNA replication

Amongst eukaryote model organisms, *S. cerevisiae* is unique in having a known conserved ACS which can be mutated to inactivate origin function (Van Houten

& Newlon 1990). Since ORC, the pre-RC proteins involved in DNA replication initiation and other features of origins (such as being nucleosome-depleted and having open chromatin) are conserved from *S. cerevisiae* to humans (Méchali 2010), this makes *S. cerevisiae* a suitable model organism in which to study DNA replication origins.

An assessment of the suitability of *S. cerevisiae* as a model organism, based on functionally classifying proteins into biological pathways and comparing these between organisms, found *S. cerevisiae* to be an appropriate model for *Homo sapiens* as the two organisms are phenotypically similar in many biological pathways (Karathia *et al.* 2011). Previously, *S. cerevisiae* has been used to model many processes related to human health including aging and neurodegenerative disorders (Miller-Fleming *et al.* 2008; Murakami & Kaeberlein 2009). It has also been used to examine signal transduction, gene expression regulation and metabolism pathways, the cell cycle and apoptosis (Brocard-Masson & Dumas 2006; Hohmann *et al.* 2007; López-Mirabal & Winther 2008; Nasheuer *et al.* 2002; Owsianowski *et al.* 2008; Rhiannon & T. Young 2009). It has been found that many genes involved in human health have yeast orthologues (Foury 1997), and over 30 % of genes identified in yeast have mammalian homologues (Botstein *et al.* 1997).

S. cerevisiae is also used widely as a eukaryotic model organism for a variety of practical reasons. The organism is non-pathogenic, and easy to acquire and then grow inexpensively in the laboratory (Sherman 2002). It has a doubling time of approximately 90 minutes, making it relatively quick compared to human cells or many other eukaryotic models (Bergman 2001). *S. cerevisiae* also has a number of genetic advantages. It can exist stably in haploid or diploid forms. In

the haploid form simple genetic modifications are possible, whilst the diploid allows independent mutations in haploid strains to be easily combined, or complementation studies (Sherman 2002). Additionally, *S. cerevisiae* cells can be easily transformed, maintain both linear and circular artificial chromosomes and plasmids and can tolerate large numbers of selective markers (Sherman 2002). Transformed DNA is only integrated into the genome by homologous recombination, meaning targeted DNA substitutions can be carried out at desired loci (Sherman 2002). Although the appearance of CRISPR-Cas9 technology (reviewed in (Jiang & Doudna 2017)) has made genetic manipulation of other models far easier, *S. cerevisiae* is unique in the number of resources available. For example, the Saccharomyces genome deletion project (SGDP) is a series of yeast strains each with one non-essential open reading frame deleted (Giaever & Nislow 2014).

1.3.1 The yeast genome

In 1996 *S. cerevisiae* became the first eukaryote to have its genome sequenced (Goffeau *et al.* 1996). In its haploid state, the *S. cerevisiae* genome is 12068 kb, split over 16 chromosomes, which range in size from 200 kb to over 1500 kb (Goffeau *et al.* 1996). There are approximately 5900 protein-coding genes, averaging one gene every 2 kb which is extremely gene dense for a eukaryote (human protein-coding gene density is 8-9 genes per Mb) (Collins *et al.* 2004; Goffeau *et al.* 1996). The high gene density observed is, in part, due to the fact that only 4 % of protein-coding genes contain introns (Goffeau *et al.* 1996). Findings from the *S. cerevisiae* genome are collated in the Saccharomyces genome database (SGD) (Cherry *et al.* 1998).

The overall genome organisation in *S. cerevisiae* is analogous to that found in metazoans (Bi 2014; Bowman & McKnight 2017; Choy *et al.* 2012; Kupiec 2014). The first examination of heterochromatin in *S. cerevisiae* was based on genetic analysis of the mating locus, *HM* (Bi 2014). Haploid yeast are one of two mating types, a or α: diploids are a/α (Hartwell *et al.* 2004). The mating type locus consists of three genetic elements, *HML*, *MAT* and *HMR* (Figure 1.4A). *HML* encodes the *α1* and *α2* genes and *HMR* encodes the *a1* and *a2* genes, but both of these regions are compacted into heterochromatin and therefore transcriptionally inactive (Bi 2014). The *MAT* locus can be *MATα* (containing *α1* and *α2* genes) or *MAT**a*** (containing *a1* and *a2* genes) (Hartwell *et al.* 2004). The *MAT* locus is transcriptionally active and this determines the cell mating type; *MATα* produces α cells whilst *MAT**a*** produces a cells (Hartwell *et al.* 2004). Heterochromatin and euchromatin (open chromatin) are defined by histone modifications such as acetylation or methylation which alter the strength of binding to other proteins and DNA (Bi 2014).

Particular histone modifications and histone variants are associated with centromeres, the site of kinetochore assembly (Choy *et al.* 2012). Correct kinetochore assembly is required for spindle attachment and subsequent chromosome segregation. The histone H3 variant, Cse4 is conserved between *S. cerevisiae* and humans (CENPA in humans), and in budding yeast and other eukaryotes H4 is hypoacetylated at centromeres (Choy *et al.* 2012). Centromeres in *S. cerevisiae* are extremely small compared to many other eukaryotes, with only 125 bp sequence required for correct chromosome segregation (Cottarel *et al.* 1989).

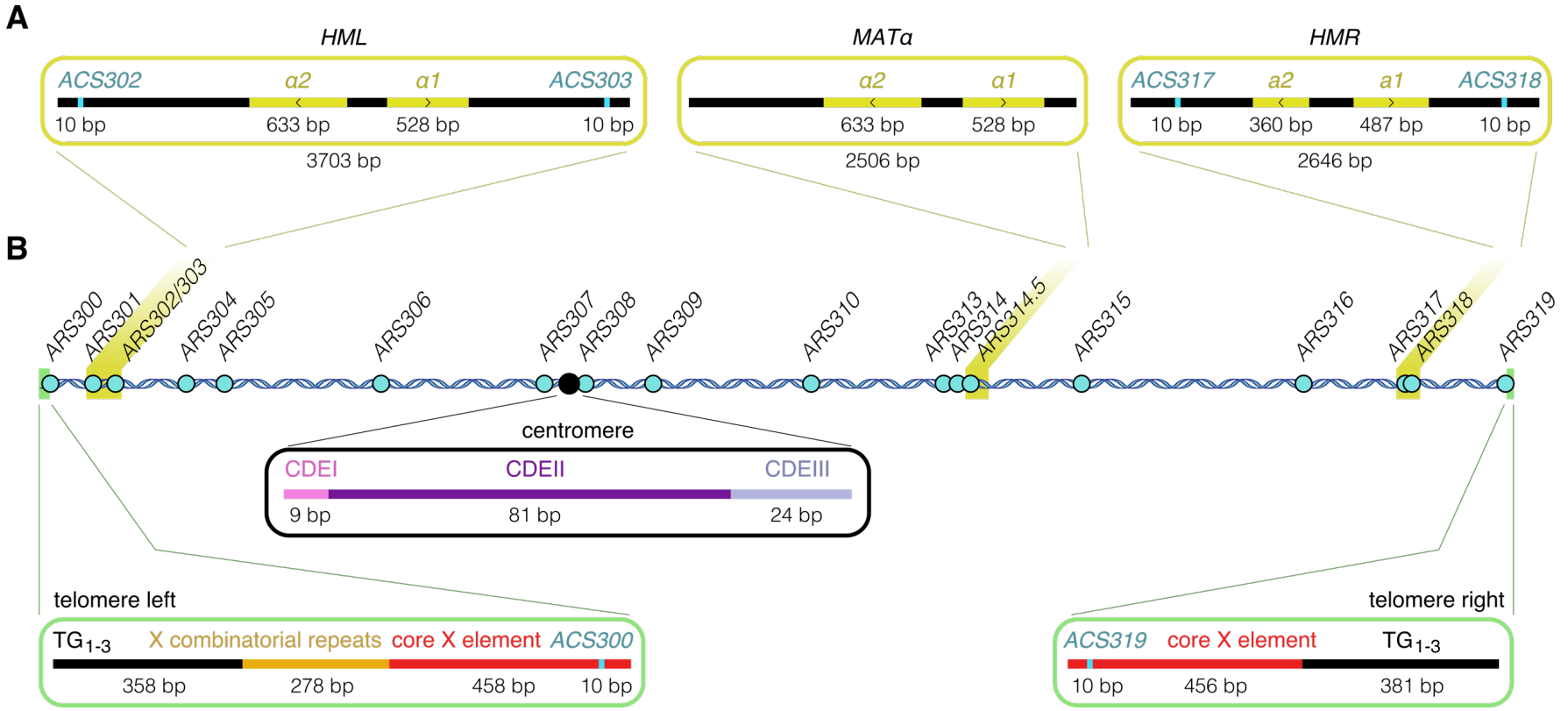


Figure 1.4* S. cerevisiae *chromosome III schematic. *Schematic representation of chromosome III genetic elements, to scale with sizes specified. A) The mating type locus,* MAT*, and the two mating type donor loci,* HML *and* HMR*. The donor loci contain mating type specific genes, and genes from either* HML *or* HMR *can be incorporated at* MAT*. B) Telomeres (green boxes), origins (blue circles) and the centromere (Black circle) are shown. Data taken from the UCSC genome browser.*

At the ends of yeast chromosomes are telomeric and subtelomeric regions (Figure 1.4B). Telomeres protect the ends of linear chromosomes from degradation, and ensure chromosome segregation (Teixeira & Gilson 2005). Yeast telomeres are composed of $(C_{1-3}A)_n$ repeats, maintained by the telomerase enzyme. This is similar to most organisms with linear chromosomes, where repeats 5 – 7 bases long which are G-rich define telomeres (McEachern & Blackburn 1994; McEachern & Hicks 1993). Subtelomeric regions contain other repeat elements; all yeast chromosome ends contain a core 'X' element which contains an ARS. Some chromosome ends (19 of 32) contain an additional Y element, which consists of 1 - 4 repeats of either a long (6.7 kb) or short (5.2 kb) Y repeat (Louis & Haber 1990, 1992). The Y element, when present, sits immediately adjacent to the functional telomere repeats, and the X element is more centromere-proximal. At some chromosome ends $(C_{1-3}A)_n$ repeats separate the X and Y elements, but this is not always the case (Louis *et al.* 1994). The X element is composed of the core X (defined as 559 bp shared by both ends of ChrIII) (Louis *et al.* 1994) and a mosaic of other repeat elements which vary between chromosome ends. Some of these repeats enable binding of proteins which aid in chromosome segregation by binding telomeres to the nuclear periphery (Hediger *et al.* 2002; Taddei & Gasser 2004).

As well as the chromosomal genetic material, the *S. cerevisiae* genome contains the mitochondrial genome (approximately 86 kb) (Foury *et al.* 1998). Some yeast cells also contain the 2 μm plasmid (6.3 kb), which carries four genes and exists in a copy number of around 60, although its absence confers no known phenotype (Futcher 1988).

1.3.2 The mitotic cell cycle in yeast

In order for eukaryotic cells to divide they must fully replicate, including their DNA, and then separate in a process called cytokinesis to generate two daughter cells (Figure 1.5) (Norbury 1992). In budding yeast, the parental cell is referred to as the mother cell, and the ellipsoidal daughter cell is formed as a bud. The yeast cell cycle, like other metazoan cell cycles, has gap stages, known as gap phase 1 (G1) and gap phase 2 (G2). G1 is the first stage of the cell cycle, when cells prepare for the DNA synthesis (S) phase and accumulate nutrients. G2 is the period during which DNA replication is completed and checked, and cells prepare for mitosis (M), which is the cell division or cytokinesis stage (Blow & Tanaka 2005). The G2 phase in *S. cerevisiae* is poorly defined; since some mitotic events occur during S phase it can be argued that S phase and M phase overlap (Forsburg & Nurse 1991; Kitamura *et al.* 2007). During *S. cerevisiae* G2, the nucleus migrates to sit between the emerging bud and the mother cell (Shaw *et al.* 1997).

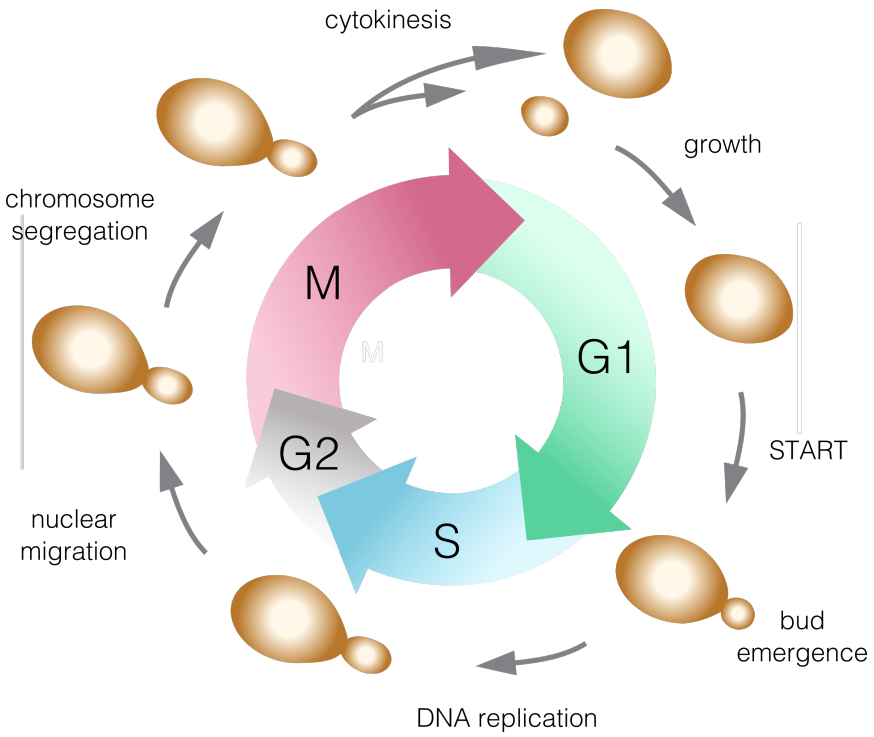


Figure 1.5 Schematic representation of the* S. cerevisiae *mitotic cell cycle. *Cell cycle phases and a representative* S. cerevisiae *cell is shown at each stage. Cells commit to a new round of mitotic cell division at START, also known as the restriction point. Bud emergence begins in G1. The genome replicates in S phase, then the nucleus migrates to sit between the mother and daughter (bud) cell in G2. Chromosome segregation occurs during mitosis and is followed by cell division (cytokinesis).*

The linking of later cell cycle events to DNA synthesis ensures that only a single round of replication occurs in each cell cycle. Replication initiates at DNA replication origins (discussed in section 1.2) during S phase. Sister chromatids are held together by cohesin until M phase (Bickel & Orr-Weaver 1996). M phase is composed of multiple stages. First, chromosomes condense to protein during prophase, then sister chromatids align along the midsection of the cell (the

metaphase plate) during metaphase. Spindles are attached to the centromeres via kinetochore attachment, and sister chromatids are separated during anaphase. During telophase, the chromosomes arrive at opposite poles of the cell. Finally, in cytokinesis the cell physically divides leaving two new cells (Pardee *et al.* 1978).

If cells exist in an environment not conducive to cell division, they can also enter a quiescent state, G0, until they are able to divide again. Environmental stresses such as oxidative or osmotic stress, or calorific limitation have been shown to extend *S. cerevisiae* lifespans by causing cells to spend more time in quiescent states (Jimenez *et al.* 2015). When growing exponentially, each cell cycle in *S. cerevisiae* lasts approximately 90 minutes, with G1, S and M phases each comprising around one third of this time (Forsburg & Nurse 1991).

1.4 Canonical DNA replication in yeast

In eukaryotes, including *Saccharomyces cerevisiae*, origin activation is a two-step process, linked to the cell cycle (for review see Nishitani & Lygerou 2002). The DNA replication origin database (OriDB) indicates hundreds of origins across the *S. cerevisiae* genome, whilst human genomes contain thousands of origins (Jeon *et al.* 2005; Méchali 2010; Siow *et al.* 2012). To ensure a single round of replication occurs per cell cycle, origin licensing (assembly of the pre-RC) occurs during G1 phase, and origin firing (activation) occurs during S phase.

1.4.1 Origin licensing (G1 phase)

Origin licensing, during which the mini-chromosome-maintenance complex (MCM, composed of Mcm2-7), which is the replicative helicase, is recruited,

occurs in G1 phase (Figure 1.6) (Blow & Hodgson 2002). The ACS is bound by the origin recognition complex (ORC) which, with licensing factors Cdc6 and Cdt1, facilitates the loading of double MCM hexamers (Diffley *et al.* 1994).

The precise origin features which recruit ORC vary between species (Méchali 2010). In *S. cerevisiae*, ORC binds DNA at ACS sites, and ATP hydrolysis by Orc1 enables recruitment of Cdc6 (Klemm & Bell 2001; Speck *et al.* 2005). After Cdc6 binding the BAH-domain at the N-terminal of Orc1 is re-orientated, which inhibits non-specific DNA binding by ORC, increasing origin specificity (Mizushima *et al.* 2000; Sun *et al.* 2012). Cdc6, alike other pre-RC proteins, is conserved in humans (Nishitani & Lygerou 2002). However, other proteins also aid in Cdc6 recruitment in metazoans (Kinoshita *et al.* 2008; Salsi *et al.* 2009).

ORC binding is limited to G1 phase. In *S. cerevisiae* cyclin-dependent kinase (CDK) binds and phosphorylates Orc6 during S and M phases, which reduces the recruitment of the replicative helicase (Chen & Bell 2011; DePamphilis 2003). In *Xenopus* the pre-RC destabilises ORC association with chromatin, and in mammals Orc1 affinity for chromatin is reduced in S phase but restored at the start of G1 (DePamphilis 2003). Cdc6 is also controlled in both *S. cerevisiae* and metazoans, either by exclusion from the nucleus or by degradation after G1 phase (Coverley *et al.* 2000; Delmolino *et al.* 2001; Jiang *et al.* 1999b; Mendez & Stillman 2000; Petersen *et al.* 1999, 2000).

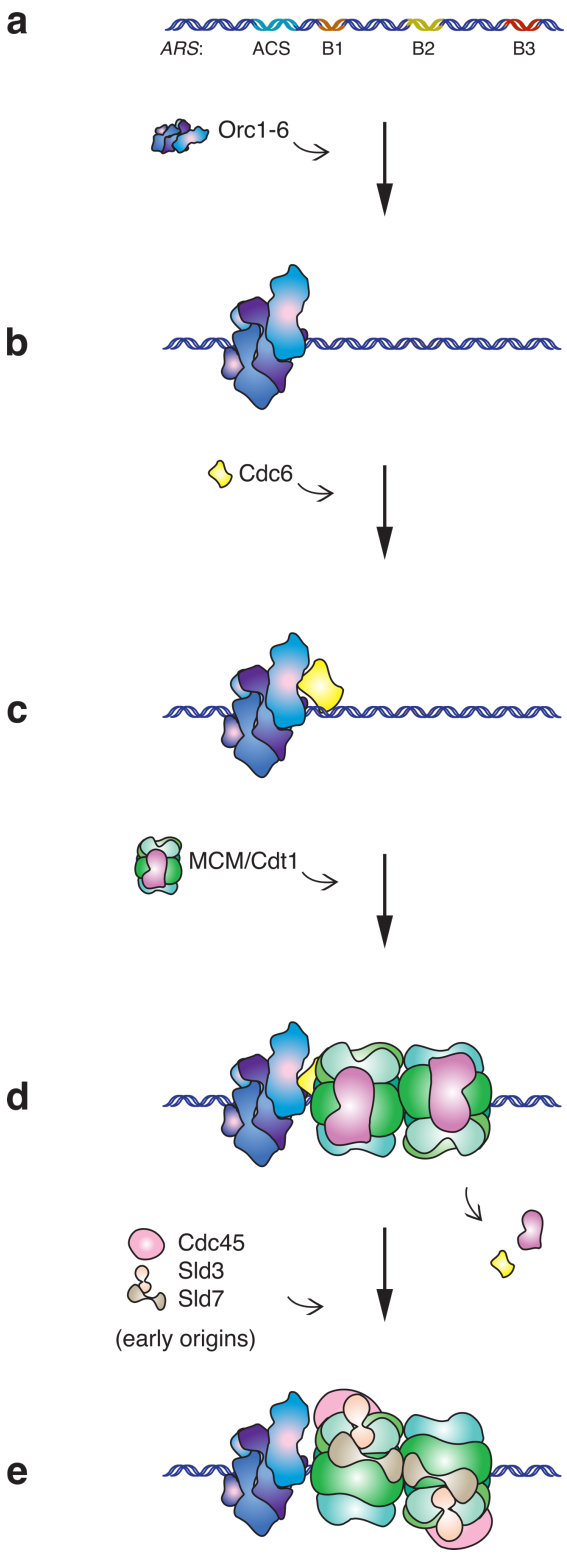


Figure 1.6 The sequence of events in canonical origin licensing. *Origin licensing occurs in G1 and involves the recognition and binding of an ACS site by origin recognition complex (ORC), Cdc6, Cdt1 and MCM helicase.*

After ORC/Cdc6 has bound the MCM replicative helicase is loaded – MCM is loaded at the B2 element in *S. cerevisiae* (Zou & Stillman 2000). The MCM complex is recruited as a heptamer, bound to Cdt1 (Remus *et al.* 2009). Cdt1 and MCM accumulation in the nucleus during G1 are interdependent (Tanaka & Diffley 2002). Orc6, which is the only ORC subunit not required for DNA binding, contains two Cdt1 binding sites (Chen *et al.* 2007; Chen & Bell 2011). In metazoans, Hbo1 acetyltransferase activates Cdt1 and Mcm9 forms a stable complex with Cdt1 to enable helicase recruitment (Iizuka *et al.* 2006; Lutzmann & Méchali 2008; Miotto & Struhl 2008). MCM has helicase activity due to the C-terminal AAA^+ ATPase domains of each MCM subunit, however it is inactive in G1 phase, requiring activation by Dbf4-dependent kinase (DDK) and CDK to function (Bochman & Schwacha 2008; Iyer *et al.* 2004).

Although MCM-Cdt1 heptamers are loaded individually, two are loaded per origin, with double hexamers connected via their N-terminal rings (Remus *et al.* 2009; Ticau *et al.* 2015). However, the precise conformation of MCMs at the origin is not known. It has been suggested that multiple MCMs may be loaded at the origin by ORC, and that early activating origins have more MCM hexamers loaded (Bowers *et al.* 2004; Das *et al.* 2015). It is also unclear whether ORC is retained during MCM loading or rebinds after MCM binds to DNA (Tsakraklides & Bell 2010).

In metazoans, geminin inhibits inappropriate MCM loading outside of G1 phase by restricting Cdt1 activity (Saxena & Dutta 2005; Tada *et al.* 2001; Wohlschlegel *et al.* 2000). It has been shown that G1 cyclins exclude Mcm4 from the nucleus prior to S phase, suggesting that the pre-RC complex assembles late in G1 (Labib *et al.* 1999). However, work in *S. pombe* has shown only assembled

MCM hexamers remain nuclear (Pasion & Forsburg 1999). In *S. cerevisiae* MCMs are actively exported to the cytoplasm during S phase and excluded from the nucleus during G2 and M phase to prevent origin licensing occurring outside of G1 (Nguyen *et al.* 2000). Further, CDK in S phase inhibits Orc2 and Orc6 subunits, preventing further helicase loading (Chen & Bell 2011).

Although DDK is recruited to MCM during late G1, it is unable to phosphorylate MCM until S phase (Lei & Tye 2001). For some early activating origins, Cdc45 in complex with Sld3 (Treslin and Ticrr in metazoans) can associate with the pre-RC in late G1 (Aparicio *et al.* 1999; Heller *et al.* 2011; Kamimura *et al.* 2001; Kumagai *et al.* 2010; Sansam *et al.* 2010). However, this interaction is only stabilised in S phase after DDK is activated (Lei & Tye 2001). This allows early activating origins to prepare to activate at the start of S phase, but prevents aberrant activation.

1.4.2 Origin firing (S phase)

During S phase, origin firing (activation) occurs. Each origin is fired at a reproducible time within S phase, and this is co-ordinated by DDK and CDK activation of the MCM complexes (Donaldson 2005). The separation of origin licensing in G1 phase and origin firing in S phase ensures each origin can be activated at most once per cell cycle. Origin firing itself is divided into two steps, regulated by DDK and CDK. It has been shown *in vitro* that there is a minimal set of factors required to initiate replication, and these are recruited in either a DDK- or CDK-dependent manner (Figure 1.7) (Yeeles *et al.* 2015).

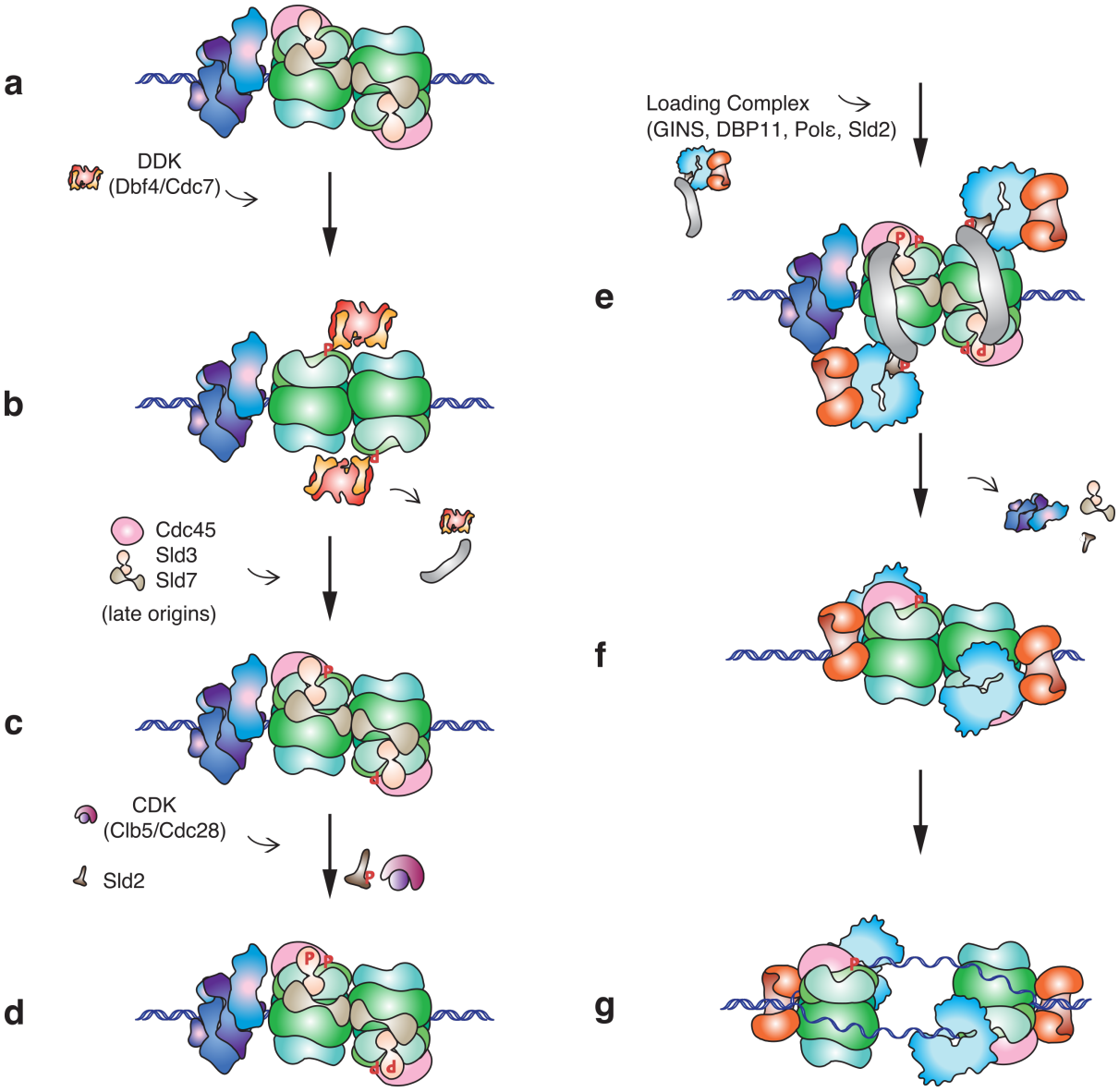


Figure 1.7 The sequence of events in canonical origin activation. *In S phase, origin firing occurs. DDK phosphorylation of MCM allows Cdc45, Sld3 and Sld7 recruitment. CDK phosphorylates Sld2 and Sld3, recruiting the loading complex (GINS, Dbp11, replicative polymerases and Sld2). ORC and other loading factors are then lost, and MCM melts (unwinds) AT-rich origin DNA, which allows DNA replication.*

DDK is composed of Dbf4 and Cdc7 subunits (Hughes *et al.* 2012b; Jares *et al.* 2000). DDK phosphorylates Mcm2, Mcm4 and Mcm6 subunits of the MCM complex at origin DNA (Lei *et al.* 1997; Masai *et al.* 2006; Sato *et al.* 1997; Sheu & Stillman 2006). This function is conserved in *Xenpous* and mammalian cells (Jiang *et al.* 1999a; Labib 2010). MCM phosphorylation leads to a structural change in the MCM complex which allows binding by Cdc45 (Labib 2010). The requirement for DDK activation is removed by a mutation in the Mcm5 subunit that mimics this conformational change (Hardy *et al.* 1997; Hoang *et al.* 2007). Other DDK targets include the polymerase α (Polα) (Nougarede *et al.* 2000).

Cdc45, Sld3 and Sld7 are all recruited in a DDK-dependent manner to phosphorylated MCM (Yeeles *et al.* 2015). Sld3 is stabilised by Sld7, and these factors bind MCM together with Cdc45 (Araki 2016). In metazoans, Cdc45 recruitment is aided by a variety of other proteins, including Treslin (the Sld3 homologue) (Balestrini *et al.* 2010; Chowdhury *et al.* 2010; Kumagai *et al.* 2011). During G1 phase, Dbf4 is targeted for degradation by the anaphase promoting complex (APC) to prevent early activation of MCM complexes (Nougarede *et al.* 2000). The APC also degrades phosphorylated Cdc45 at the end of S phase, inhibiting it outside of S phase (Pollok & Grosse 2007).

CDK activity is required after DDK to enable origin activation (Yeeles *et al.* 2015). In *S. cerevisiae* two cyclins can associate with CDK during S phase. Clb6 and Clb5 are both synthesised late in G1, and are closely related (Jackson *et al.* 2006). However, Clb6 is present in early S phase, and degraded rapidly at onset of S phase, therefore only early origins are activated by CDK-Clb6. Conversely, late origins are activated by CDK-Clb5 (Jackson *et al.* 2006). DNA replication is delayed in a Δ*CLB5* mutant, whereas DNA replication is not replayed in a Δ*CLB6*

mutant (Jackson *et al.* 2006). The difference in Clb5/Clb6 function is due to the regulation of their stability – Clb5 is present throughout S phase whilst Clb6 is not (Jackson *et al.* 2006). CDK phosphorylates Sld3 at the origin site and Sld2, which is a component of the pre-loading complex (pre-LC) (Tanaka *et al.* 2007).

The pre-LC consists of Sld2, the leading strand polymerase (Polε), Dbp11 and the GINS complex (Muramatsu *et al.* 2010). These proteins associate in the pre-LC prior to chromatin association. In *S. cerevisiae*, Sld2 is unusual in having 11 CDK phosphorylation motifs, most of which modify the availability of a single motif (Tak *et al.* 2006). Phosphorylation of Sld2 by CDK at this key motif facilitates binding of Sld2 to Dpb11 (Masumoto *et al.* 2002; Tak *et al.* 2006). This exquisite control over Sld2 phosphorylation is thought to prevent premature activation of origins by CDK. The GINS complex is comprised of four proteins: Sld5, Psf1, Psf2 and Psf3 (MacNeill 2010; Takayama *et al.* 2003). Polε associates with Dbp11-Sld2 and Psf1, forming the pre-LC complex (Edwards *et al.* 2003; Takayama *et al.* 2003).

The human Sld3 homologue (Treslin/Ticrr) is phosphorylated by CDK at conserved phosphorylation sites, and associates with TopBP1 in a CDK-dependent manner (Boos *et al.* 2011; Kumagai *et al.* 2011). The human homologue of Sld2 (RecQL4) is also essential for replication initiation, but it is not known if it is phosphorylated (Matsuno *et al.* 2006; Sangrithi *et al.* 2005). In human cells Mcm10 to binds to MCM, and recruits RecQL4 and Polα, along with other replication initiation factors (Xu *et al.* 2009; Zhu *et al.* 2007).

Association of the pre-LC with the pre-RC brings GINS into contact with the MCM and Cdc45. GINS is essential for replication elongation but Sld3 is not (Kanemaki & Labib 2006). GINS competes for MCM and Cdc45 binding,

displacing Sld3 *in vivo* (Bruck & Kaplan 2011b). *In vitro*, MCM Cdc45, Sld3 and GINS are able to form both Cdc45/MCM/Sld3 (CMS) and Cdc45/MCM/GINS (CMG) complexes (Bruck & Kaplan 2011a). *In vivo*, the CMG proteins are tightly associated; Cdc45 and GINS activate the MCM helicase, increasing the rate of ATP hydrolysis two-fold (Ilves *et al.* 2010). In order to form, the CMG requires both Dbp11 and phosphorylated Sld3 present at the origin; however, a phosphorylation mimic Sld2 mutant combined with an Sld3-Dbp11 fusion protein can bypass this requirement for CDK (Zegerman & Diffley 2007).

Although GINS has many roles, its main function in the CMG is to hold Cdc45 and MCM together to function as the active replicative helicase (Gambus *et al.* 2006; Moyer *et al.* 2006). Structurally, the MCM hexamer forms a gapped ring which is completed by Cdc45 and GINS (Costa *et al.* 2011). The two CMGs melt DNA at the origin and progress, bi-directionally, to replicate DNA (Yeeles *et al.* 2015). Importantly, the CMG progresses N-terminally, meaning the two CMG pass one another within the origin, synthesising the DNA between the double hexamer first (Douglas *et al.* 2018). The CMG is a conserved structure, required throughout DNA replication elongation (Ilves *et al.* 2010; Pacek & Walter 2004).

1.4.3 Replication elongation

As predicted by Watson and Crick (Watson & Crick 1953), DNA replication has been found to be semiconservative across all domains of life (Burgers & Kunkel 2017). Broadly speaking, as the replication fork progresses the leading strand is synthesized 5’ to 3’ continuously while the lagging strand is synthesised discontinuously in small 5’ to 3’ fragments, known as Okazaki fragments (Okazaki *et al.* 1968). However, errors in replication can lead to discontinuity in the leading

strand too (McElhinny *et al.* 2010b; Tye *et al.* 1977). DNA synthesis is performed by three B-family polymerases, polymerase ε (Polε), polymerase δ (Polδ) and polymerase α (Polα) (for reviews see Johansson and Dixon, 2013; Makarova, Krupovic and Koonin, 2014). It was originally proposed that all three polymerases were essential and had separate functions in DNA replication (Morrison *et al.* 1990).

Ribonucleotide incorporation studies have supported this theory. In *S. cerevisiae*, one ribonucleotide is accidentally incorporated by the Polα, Polδ and Polε every 625, 5000 and 1250 nucleotides respectively and these rates are similar in mammals (Clausen *et al.* 2013; Göksenin *et al.* 2012; Nick McElhinny *et al.* 2010). Deletion of the RNaseH2 gene (which repairs ribonucleotide incorporation) (Eder & Walder 1991; Sparks *et al.* 2012) allows accumulation of ribonucleotides in the context of less nucleotide-specific polymerase mutants (Lujan *et al.* 2013; McElhinny *et al.* 2010b; Miyabe *et al.* 2011). These studies have shown mutant Polε leads to ribonucleotide incorporation on the leading strand (McElhinny *et al.* 2010b; Miyabe *et al.* 2011), and mutant Polα and Polδ lead to ribonucleotide incorporation on the lagging strand (Figure 1.8) (Lujan *et al.* 2013).

In contradiction to this idea, it has been shown in budding and fission yeast that the Polε is not essential for cell viability (Dua *et al.* 1999; Feng & D'Urso 2001; Kesti *et al.* 1999; Ohya *et al.* 2002). It has been suggested that Polδ may be the major replicative polymerase on both strands, and Polδ is capable of initiating DNA synthesis on the leading strand (Johnson *et al.* 2015; Yeeles *et al.* 2016). However, Polε is exclusively able to engage the CMG for maximum replication processivity (Yeeles *et al.* 2016). Further, mutant studies looking at

nucleotide mis-matches resulting from ribonucleotide incorporation support a model where Polε predominantly replicates the leading strand and Polα and Polδ predominantly replicate the lagging strand (Lujan *et al.* 2012, 2013; McElhinny *et al.* 2008, 2010a; Miller-Fleming *et al.* 2008; Pursell *et al.* 2007).

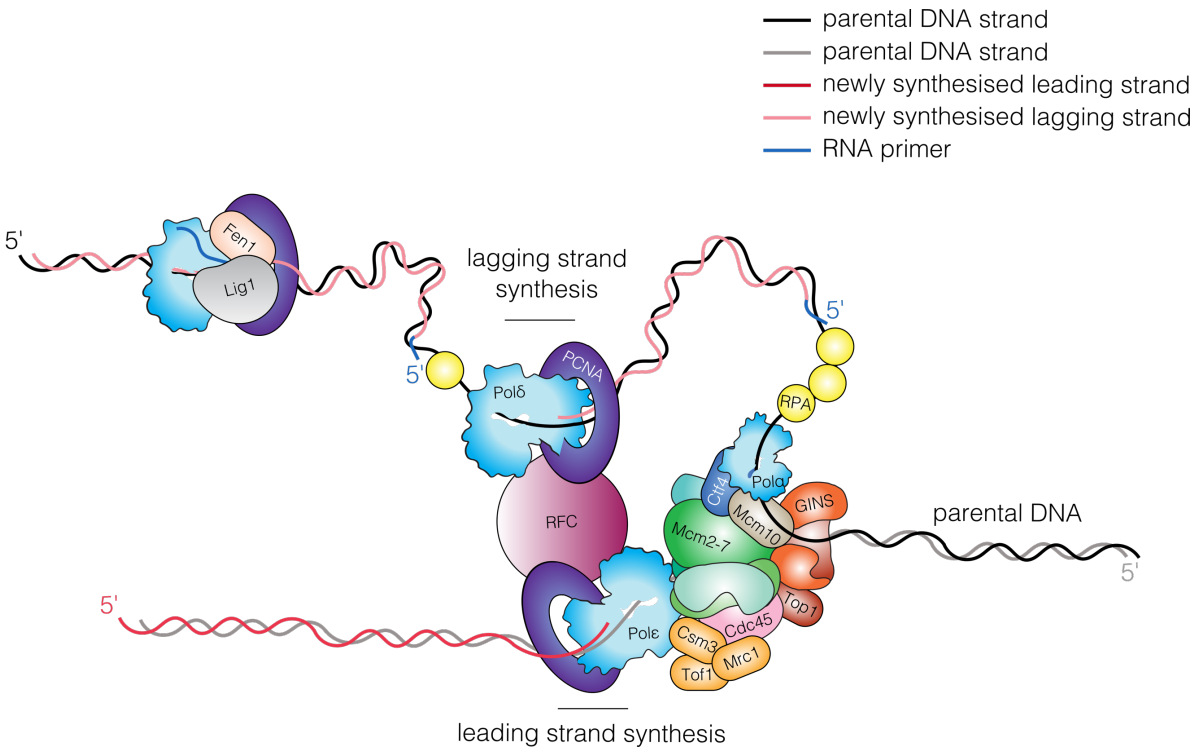


***Figure 1.8 Schematic of an ongoing replication fork in* S cerevisiae.** *Proteins discussed in this chapter associated at the replication fork are labelled; other replication fork proteins are excluded for clarity. Parental DNA is unwound by the CMG replicative helicase. Polε continuously replicates the leading strand (red), while Polδ discontinuously replicates the lagging strand (pink) in Okazaki fragments. Polα primes new Okazaki fragments on the lagging strand (blue). Fen1 and DNA ligase 1 remove 5' RNA primer flaps at the end of Okazaki fragment synthesis. Single stranded DNA is bound by RPA to prevent intra-strand annealing and DNA nuclease function.*

When the CMG is activated at the origin, it switches from encircling double-stranded DNA (dsDNA) to encircling single-stranded DNA (ssDNA) of the leading strand (Fu *et al.* 2011; Muramatsu *et al.* 2010; Sengupta *et al.* 2013). Interactions between Mcm2 and Mcm5 are important for this conformational change (Bruck & Kaplan 2015; Samel *et al.* 2014). The Polε small subunit also encircles DNA and interacts with the CMG, allowing processive 5' to 3' replication of the leading strand (Hogg *et al.* 2014). The 'trombone' model for DNA replication involves the lagging strand of DNA being bent backwards, allowing the replication fork to proceed 5' to 3' along the DNA (Sinha *et al.* 1980). This model is supported by electron microscopy and single molecule studies (Chastain *et al.* 2003; Hamdan *et al.* 2009). In yeast, Ctf4 bridges the two polymerases at the replication fork (Simon *et al.* 2014; Villa *et al.* 2016). Ctf4 is able to bind GINS and Polε on the leading strand, and Polα on the lagging strand (Bermudez *et al.* 2010; Gambus *et al.* 2009). In humans, Ctf4 binds to Polδ too (Bermudez *et al.* 2010). During elongation, dNTP levels limit DNA replication (Dieckman *et al.* 2010; Nick McElhinny *et al.* 2010) to approximately 50 nucleotides per second (Raghuraman *et al.* 2001).

Polα primes replication on both leading and lagging strands, but is more present on the lagging strand due to its need for frequent re-priming of Okazaki fragments (Pellegrini 2012). A short ribonucleotide primer of approximately 10 nucleotides is synthesised by the primase subunit of Polα, followed by 20 – 30 nucleotides of lagging strand synthesis by the Polα polymerase subunit (Bullock *et al.* 1991; Nethanel *et al.* 1988). Replication factor C (RFC), which is conserved in eukaryotes (Cullmann *et al.* 1995; Furukawa *et al.* 2003), binds to the 3' end of

the DNA-RNA primer (Furukawa *et al.* 2003). RFC then loads proliferating cell nuclear antigen (PCNA) at the replication fork (Majka & Burgers 2004). Both Polε and Polδ are bound to PCNA at the replication fork, which also stimulates processivity, particularly for Polδ (Chilkova *et al.* 2007). The switch from Polα to Polδ may be caused by RFC displacing Polα primase (Furukawa *et al.* 2003; Tsurimoto & Stillman 1991; Yuzhakov *et al.* 1999), or due to the conformation of dsDNA becoming less favourable to Polα as replication progresses (Perera *et al.* 2013).

FEN1 is associated with the PCNA homotrimer on the lagging strand, in a 'toolbelt' model (Stodola & Burgers 2016). FEN1 cleaves 5' flaps generated when the Polδ reaches the 5' end of the previous Okazaki fragment, removing the ribonucleotide primer created by primase/Polα (Balakrishnan & Bambara 2013; Tsutakawa *et al.* 2014). The Polδ disassembles quickly to prevent the flaps being enlarged (Burgers 2009). When large flaps are created, Dna2 can cleave them, but is less precise (Kang *et al.* 2010; Wanrooij & Burgers 2015) and FEN1 is thought to function after Dna2. The Okazaki fragment maturation machinery comprises Polδ, FEN1 and the DNA ligase (Cdc9 in *S. cerevisiae*), which seals nicks on the lagging strand (Vijayakumar *et al.* 2007). Although the ligase can bind to PCNA (Vijayakumar *et al.* 2007), it does not appear to remain associated with the replication fork (Ayyagari *et al.* 2003). Replication protein A (RPA) is also present on the lagging strand. RPA coats ssDNA exposed during Okazaki fragment synthesis (Longhese *et al.* 1994).

Other proteins also associate with the replication fork. Tof1-Csm-Mrc1 form the replication pausing complex, which is present in the replisome (Calzada *et al.* 2005). Tof1-Csm-Mrc1 associate with Cdc45 on the leading strand of DNA,

and can activate the intra-S phase checkpoint if the replication fork stalls (Alcasabas *et al.* 2001; Kumagai & Dunphy 2000). This complex is conserved in mammals as the Tim1-Csm-Claspin complex (Katou *et al.* 2003; Prakash *et al.* 2005; Takahashi *et al.* 2005). When DNA lesions are encountered, specialised DNA translesion polymerases are recruited by PCNA to temporarily synthesis DNA (Burgers *et al.* 2001; Prakash *et al.* 2005; Takahashi *et al.* 2005). These polymerases lack proof-reading, so are switched out in favour of Polε or Polδ once the lesion is bypassed. DNA topoisomerases I and II (TopI and TopII) function to prevent supercoiling of DNA ahead of (TopI) and behind (TopII) the replication fork (Bermejo *et al.* 2007). TopII functions to remove catenation of replicated DNA behind the replication fork (Baxter & Diffley 2008; Lucas *et al.* 2001). Finally, histone chaperones such as FACTs (facilitates chromatin transcription) Spt16 and Pob3 aid in chromatin re-modelling around DNA replication (Alabert & Groth 2012). In total, more than 50 different proteins may associate with the replication fork during elongation (Blow *et al.* 2011).

1.4.4 Replication termination

In mammalian cells, replication termination occurs approximately 50,000 times per S phase (Berezney *et al.* 2000). Some yeast termination sites are reproducible, leading to the suggestion of a genetically determined termination sequence (Fachinetti *et al.* 2010). However, it has now been shown that termination predominantly occurs stochastically between two origins (fork merger zones), and when patterns in origin initiation are changed predictable shifts in the location of fork merger zones occur (Greenfeder & Newlon 1992; McGuffee *et al.*

2013). A select few site-specific termination zones do exist in *S. cerevisiae*, and are discussed at the end of this section.

1.4.4.1 Fork merger zones

Five things occur when two replication forks converge (Dewar & Walter 2017). The first is torsional stress; supercoiled DNA builds in the parental DNA in front of the replication fork. However when two replication forks are closer than approximately 150 bp the parental DNA is unable to twist due to the stiffness of DNA (Vafabakhsh & Ha 2012). Despite supercoiling occurring in *Xenopus* egg extracts, DNA replication elongation does not slow down in the 500 bp approaching convergence of replication forks (Dewar *et al.* 2015), indicating that supercoiling does not limit DNA replication termination.

Instead of supercoiling, the entire replication fork can be rotated, forming pre-catenanes (Postow *et al.* 2001). Pre-catenanes can only be resolved by TopII (Postow 2001). In both *S. cerevisiae* and *Xenopus* egg extracts, convergence and completion of synthesis was not dependent on TopII (Baxter & Diffley 2008; Dewar *et al.* 2015; DiNardo *et al.* 1984). This is in contrast to prokaryotic termination of replication (reviewed in Dewar and Walter, 2017). Catenanes are formed from pre-catenanes after the final turn of parental DNA has been replicated (Ullsperger *et al.* 1995). The replicon size had no impact on the number of catenanes formed in *S. cerevisiae*, indicating that pre-catenanes are primarily formed during termination (Lucas *et al.* 2001; Schalbetter *et al.* 2015; Ullsperger *et al.* 1995).

The second thing that occurs in converging replisomes is they must encounter one another. Since the CMG encircles the leading strand only (Fu *et*

al. 2011), it has been suggested that the two replisomes may bypass one another without clashing (Dewar & Walter 2017). This theory is supported by the weak interaction between the leading and lagging strand complexes in the replisome (section 1.4.3). Further, it has recently been shown that the CMG pass one another within the origin when initiating replication (Douglas *et al.* 2018). However, CMG stalling is observed when there are protein-DNA interactions on the lagging strand (Duxin *et al.* 2014), therefore the question of whether converging CMG can clash is not fully resolved.

Thirdly, the replisome must be removed from the termination site to prevent re-replication. *In vitro* systems that recapitulate replication initiation and elongation, demonstrate inefficient termination (Devbhandari *et al.* 2017). Contrastingly, *Xenopus* egg extracts demonstrate efficient termination, suggesting there is an active process occurring (Dewar *et al.* 2015; Moreno *et al.* 2014). It has recently been shown that replication disassembly in eukaryotes is a carefully controlled active process.

During termination both in yeast and in frogs, Mcm7 is ubiquitylated (Maric *et al.* 2014; Moreno *et al.* 2014). Ubiquitylation is performed by an E3-ubiquitin ligase: SCF^{Dia2} ([Skp, Cullin, F-box containing complex] associated with [Digs into agar 2]) in yeast and $CRL2^{Lrr1}$ ([Cullin RING Ligase 2] associated with [Leucine Rich Repeats 1]) in frogs and worms (Dewar *et al.* 2017; Sonneville *et al.* 2017). In yeast, SCF^{Dia2} associates with the traversing replisome, whilst in frogs, $CRL2^{Lrr1}$ is recruited to the replisome only during termination (Dewar *et al.* 2017; Gambus *et al.* 2006). It should be noted that the experiments in *Xenopus* extracts examined replication termination on a DNA plasmid, which may affect whether the $CRL2^{Lrr1}$ is able to be recruited to the replisome prior to termination (Dewar *et*

al. 2017). Mcm7 ubiquitylation results in CMG removal from DNA by the ATPase p97 (Franz *et al.* 2011; Maric *et al.* 2014; Moreno *et al.* 2014). The CMG does not appear to be degraded in frogs, and may be recycled in future cell cycles (Fullbright *et al.* 2016).

Removal of the CMG from termination sites is crucial for eukaryotes. A lack of Dia2 leads to chromosome instability in yeast (Blake *et al.* 2006; Koepp *et al.* 2006; Pan *et al.* 2006). In order to prevent unstable chromosomes, worms have a back-up pathway to remove the CMG during prophase which is dependent on UBXN-3 (a co-factor of the ATPase Cdc-48) (Sonneville *et al.* 2017). Inactivation of both pathways is synthetic lethal. Further study is required to investigate the CMG removal pathway which occurs in prophase in frogs.

Unloading occurs after the synthesis of DNA is completed (Dewar *et al.* 2015). When the CMG reaches the 5' end of the last Okazaki fragment synthesised by the converging replisome it is able to translocate over dsDNA (Kang *et al.* 2013; Kaplan *et al.* 2003). This suggests ubiquitination of Mcm7 occurs when the CMG encircles dsDNA (Dewar & Walter 2017). In support of this theory, CMG can be unloaded in the G1 phase by recruited $CRL2^{Lrr1}$(frogs) or activated SCF^{Dia2} (yeast) (Maric *et al.* 2014). In order to distinguish between pre-RC loaded MCM and CMG complexes, the E3-ubiquitin ligases may be able to recognise Cdc45 or GINS (Dewar & Walter 2017). A process that unloads the CMG from dsDNA is appealing as it would have the potential to distinguish between ongoing replication forks and those that are ready to terminate. However, the CMG is also unloaded from ssDNA when the replisome encounters inter-strand crosslinks (Fu *et al.* 2011; Long *et al.* 2014). A different ubiquitin

ligase is required for this process, indicating there is still much to learn about active CMG unloading during replication termination.

The rest of the replisome is presumed to dissociate passively after CMG removal. Polε and CMG have similar dissociation kinetics, and when CMG unloading is blocked the rest of the replisome remains on chromatin, adding support to this hypothesis (Dewar *et al.* 2015, 2017). PCNA is constantly turned over at the replisome by ELG1, and the Okazaki maturation machinery Polδ, FEN1 and the Cdc9 ligase is frequently turned over during elongation (Dungrawala *et al.* 2015; Kubota *et al.* 2013).

The fourth important aspect of converging replication forks is the replication of the ssDNA remaining between the two replisomes (gap filling). This has already been touched upon above. In eukaryotes where the CMG are able to pass one another, it appears large stretches of ssDNA are not left exposed. DNA synthesis of the 'gap' is completed even when unloading of the CMG is blocked, adding weight to the idea that unloading of the CMG occurs after synthesis is fully complete (Dewar *et al.* 2015, 2017). It is important for the CMG to translocate to allow FEN1 and DNA ligase access to the DNA. Rapid ligation of the synthesised gap and the nearest Okazaki fragment is observed in *Xenopus* (Dewar *et al.* 2015). It is also important for the CMG to be removed rapidly after it translocates onto dsDNA in order to prevent re-replication from downstream Okazaki fragment 5' flaps. Loss of the E3-ubiguitin ligase $CRL2^{Lrr1}$ in worms leads to re-replication (Merlet *et al.* 2010).

The final, fifth, aspect of converging forks is the resolution of catenanes left after replication prior to mitosis. Yeast has a decatenation process distinct from that in *E. coli* (Baxter *et al.* 2011; Zechiedrich *et al.* 1997). Decatenation in

yeast is dependent on condensin loading and spindle formation, indicating that it occurs after the termination of replication forks (Baxter *et al.* 2011).

1.4.4.2 Site-specific termination

Site-specific replication termination is rare in *S. cerevisiae*, but at least two forms of it exist (Bastia & Zaman 2014; Dalgaard *et al.* 2009). The first occurs at replication fork barriers (RFBs), which stall replication forks from one direction (Fachinetti *et al.* 2010). The most common RFB in budding yeast is found in the ribosomal DNA (rDNA) locus repeats (Figure 1.9). At each rDNA repeat a RFB exists which consists of a *Ter* sequence, bound by the fork blocking protein Fob1 (Bairwa *et al.* 2010). Like the *E. coli* terminator, the function of *Ter* in rDNA prevents transcription-replication head on collisions which might otherwise lead to genome instability (Hamperl & Cimprich 2016). Although fork arrest at RFBs is well defined, how the replisome is dissociated is not (Jaiswal *et al.* 2016; Mohanty *et al.* 2006). In *Xenopus* stalled replication forks can restart (Dewar *et al.* 2015), so it is possible that once the converging fork arrives at the termination site the replisomes pass one another and are unloaded as described in section 1.4.4.1. However, there may be a specialised role for the helicase rDNA recombination mutation protein 3 (Rrm3) as when Rrm3 is inactivated, replisomes accumulate at RFBs (Ivessa *et al.* 2000).

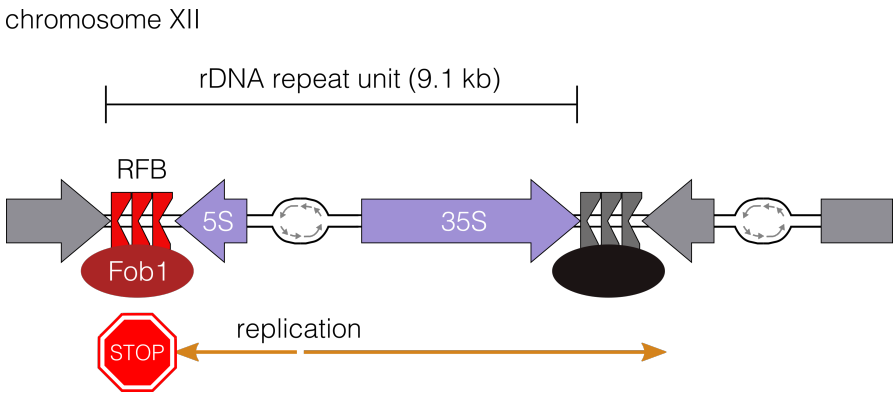


***Figure 1.9 Schematic of the rDNA replication fork barrier in* S. cerevisiae.** *The rDNA array* in S. cerevisiae *is 9.1 kb and consists of two transcription units, 35S and 5S; the origin is positioned between these transcription units. The replication fork barrier (RFB) stops replication uni-directionally. The RFB prevents collisions between transcriptional and replication machinery in the 35S coding sequence. The RFB is bound by Fob1.*

The second known type of site-specific replication termination type known occurs at telomeres (Martínez & Blasco 2015). The leading strand is replicated to within a few base pairs of the telomere end (Chow *et al.* 2012; Soudet *et al.* 2014; Wu *et al.* 2012a). This mimics the result of using purified Polε *in vitro* and suggests that the CMG slides off the end of the chromosome (Hogg *et al.* 2014; Sfeir *et al.* 2005). As a result of this, the Polα which is tethered to the CMG is lost, and cannot re-prime the lagging strand. This results in unreplicated DNA at the telomeric end of the lagging strand approximately equal to the length of an Okazaki fragment (70 – 100 bp) (Chow *et al.* 2012).

1.5 Control of the cell cycle in yeast

1.5.1 Kinases

Control of the cell cycle is largely attributable to the sequential activation and inactivation of protein kinases, most notably the CDKs (Suryadinata *et al.* 2010). Cell cycle control by CDKs is conserved between yeast and metazoan cells (Sherr & Roberts 2004). In budding yeast, CDK consists of the kinase Cdc28 and various cyclins which function at different cell cycle stages to regulate the activity of CDK (Morgan 1997). There are multiple levels of regulation to ensure that only the correct kinases are active (Figure 1.10).

As cells exit mitosis, Cdc14 phosphatase activates a component of the APC, which leads to degradation of non-G1 cyclins (Clb proteins) (Wolfe *et al.* 2006). This promotes cells to enter G1 phase, which is characterised by low CDK activity. Cdc14 further acts to activate a transcription factor, Swi5, by dephosphorylating it. Swi5 activates transcription of Sic1, a cyclin dependant kinase inhibitor (CKI). Collectively, the low CDK state of G1 enables the binding of DNA replication origins by initiator proteins ORC and the pre-RC (origin licensing) but does not allow replication initiation to occur. A G1 cyclin, Cln3, becomes available when environmental cues indicate that cell division is possible. Cln3 binds to CDK and triggers entry into a new cell cycle. If Cln3 levels are too low, cells enter the G0 state by upregulating the kinase Rim15 (Jimenez *et al.* 2015). However, if Cln3 levels are sufficient, CDK levels increase, and phosphorylate both the APC and Sic1, inactivating them. This promotes entry into S phase by allowing accumulation of newly synthesised genes.

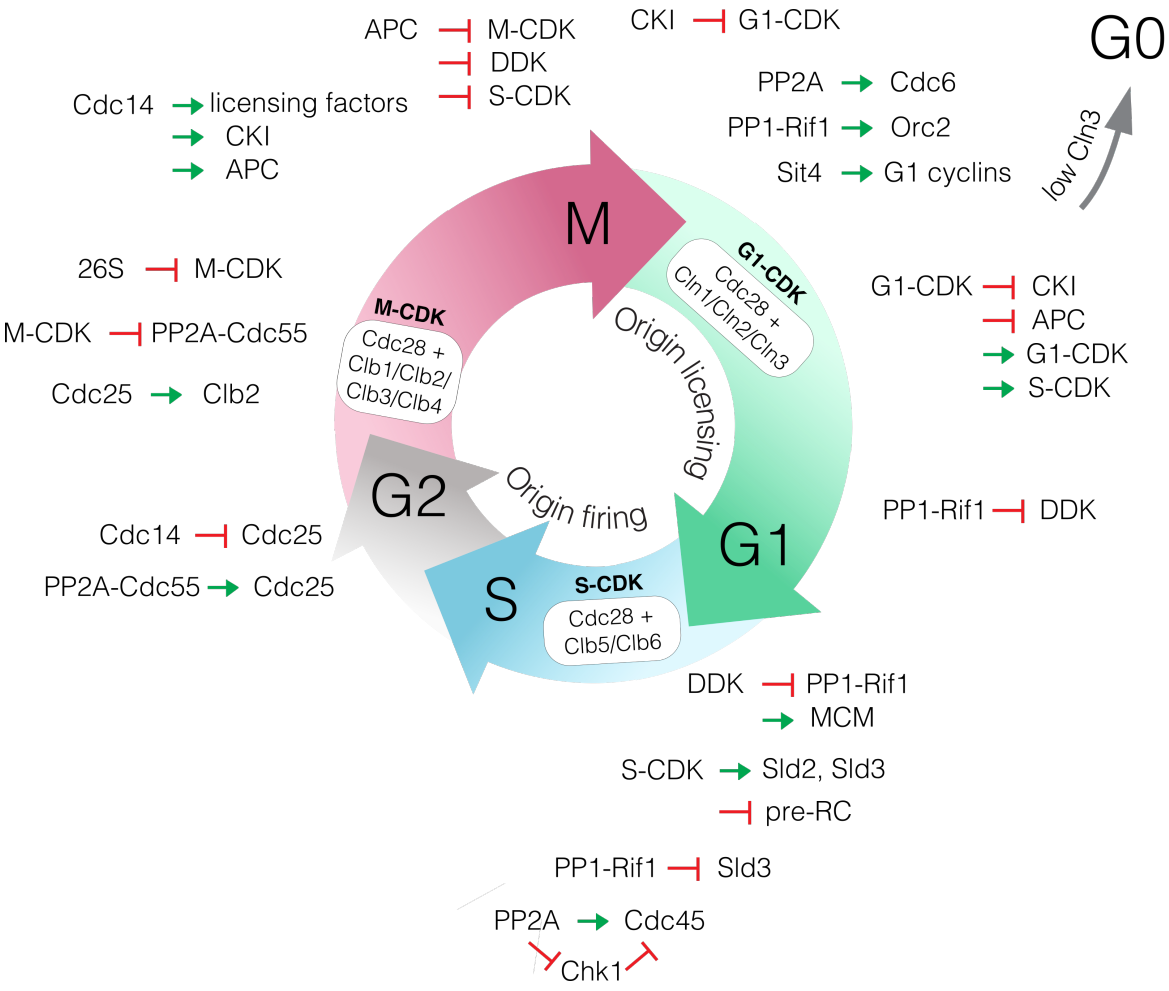


***Figure 1.10 Summary of discussed kinases and phosphatases which contribute to cell cycle control in* S. cerevisiae.** *Throughout the cell cycle different kinases and phosphatases act. Each kinase/phosphatase promotes (green arrow) or restricts (red lines) activity of other substrates. The CDK composition changes throughout the cell cycle. G1-, S- and M-CDK components are listed (white boxes). Low Cln3 levels in G1 phase leads to senescence (G0) prior to the START checkpoint.*

High CDK levels act to prevent further origin licensing, by phosphorylating ORC and other pre-RC components. S phase cyclins (Clb5 and Clb6) also bind and sequester some pre-RC components (Morgan 1997). Furthermore,

transcription of replication factor genes is promoted. Once APC levels drop, the DDK, which is an APC substrate, is stabilised, allowing it to phosphorylate the pre-RC complex and promote origin activation. CDK phosphorylation of DNA replication proteins is also required to activate the DNA unwinding helicase prior to replication initiation.

During G2/M phase, pre-RC components are sequestered by the M phase cyclin Clb2 (Honey & Futcher 2007; Loog & Morgan 2005). As mitosis progresses, the 26S proteasome degrades Clb2 (Goh *et al.* 2000). Pre-RC components are then released - enabling a new round of origin licensing and DNA replication.

Careful control of CDKs allows the function of checkpoints, which govern progression through the cell cycle. The first of these occurs in the G1 phase, at the point that cells commit to entering the cell cycle (discussed above). This checkpoint is known as START in yeast, or the restriction point in metazoans (Johnson & Skotheim 2013). The second checkpoint occurs at the G2/M, in metazoans (Stark & Taylor 2004). However, two pathways which monitor DNA damage and DNA replication stress also feed into a checkpoint in yeast which prevents chromosome segregation in the presence of significant DNA damage (Gardner *et al.* 1999). Finally, a third checkpoint occurs at the metaphase-to-anaphase transition. This ensures sister chromatids are aligned on the metaphase plate and connected to mitotic spindles via kinetochores. The metaphase-to-anaphase transition occurs when the APC is activated (Cohen-Fix & Koshland 1997).

1.5.2 Phosphatases

Whilst the role of kinases in cell cycle control has been well defined, the role of phosphatases was initially considered to be less important. This is due, in part, to there being fewer phosphatases than kinases, with those that are known having little specificity (Virshup & Shenolikar 2009). There are only around 40 serine/threonine specific phosphatases, which is 10 % of the number of kinases (Stark *et al.* 2015). The serine/threonine phosphatases include the phosphoprotein phosphatases (PPPs). Despite the low number of PPP phosphatases, they have high specificity for substrates via regulatory subunits, which direct the phosphatase catalytic subunit (Stark *et al.* 2015). The PPPs include the protein phosphatase 1 (PP1) and protein phosphatase 2A (PP2A) phosphatases. PP1 is regulated by more than 90 different subunits in humans, whilst PP2A has only 13 regulatory subunits in humans, and three in yeast (Blow & Dutta 2005). It is now known phosphatases also play a key role in controlling the cell cycle (Figure 1.10).

At mitotic exit, both PP1 and PP2A function to reverse CDK phosphorylation events ready for G1. In mammals, PP1 and PP2A each promote recruitment of mitotic exit regulatory subunits for the other, leading to a phosphatase relay and rapid activation of PP1 and PP2A activity (Grallert *et al.* 2015). In budding yeast, the key mitotic exit phosphatase, Cdc14, regulates PP1 and PP2A activity (Ramos *et al.* 2019). This helps to re-establish pre-RC competence by activating licensing factors. For example, in human cells ORC2, is dephosphorylated by PP1 (Lee *et al.* 2014); dephosphorylation of Orc2 allows the process of origin licensing to occur. Cdc6 is dephosphorylated by PP2A (Davis *et al.* 2008) and when this function of PP2A is inhibited, cells arrest in G1,

unable to progress into S phase (Davis *et al.* 2008). PP2A-like phosphatases are also important in cell cycle entry; in the absence of PP2A-ike phosphatase Sit4, expression levels of G1 cyclins are reduced (Fernandez-Sarabia *et al.* 1992).

Phosphatases also function during S phase. The Rif1 subunit is able to recruit PP1 to MCM subunits (e.g. Mcm4) and dephosphorylate them (Hiraga *et al.* 2014). This counteracts DDK activity in G1 phase. A greater DDK concentration is therefore required to promote origin firing, since the MCM phosphorylation rate must exceed its dephosphorylation rate. Conversely, DDK can bind directly to Rif1 and inhibit its interaction with PP1 (Alver *et al.* 2017; Poh *et al.* 2014). Therefore, as DDK levels increase during S phase, MCM phosphorylation is promoted and dephosphorylation is inhibited. The resulting feedback loop allows for a rapid switch from low MCM phosphorylation in G1 to high MCM phosphorylation in S phase. Having antagonistic phosphorylation and dephosphorylation of MCM gives a tighter control over origin activation (Poh *et al.* 2014; Yamazaki *et al.* 2012). Rif1-PP1 involvement in DNA replication control appears to be conserved throughout eukaryotes, as both *Xenopus* egg extract and HeLa cell studies support the findings in yeast (Stark *et al.* 2015). Post-DDK activity, in yeast, Rif1-PP1 may also antagonise CDK phosphorylation (Mattarocci *et al.* 2014). In *RIF1* null yeast strains phosphorylation of Sld3, but not Sld2, is increased (Hiraga *et al.* 2016). Conversely, PP2A is required for Cdc45 loading onto chromatin (Yan *et al.* 2000), suggesting it promotes origin activation.

PP2A contributes to regulating cell cycle checkpoint activation. In *Xenopus* egg extracts, PP2A antagonises the DNA damage pathways in response to DSBs (Petersen *et al.*, 2006). By dephosphorylating Cdc45, PP2A opposes the inhibitory action of Chk1 checkpoint protein, allowing replication to continue

(Petersen *et al.*, 2006): In higher eukaryotes, phosphorylation of Chk1 is antagonised by PP2A (Leung-Pineda *et al.* 2006). It is also suggested that PP2A inhibits Chk1 directly, although an interaction has not been shown (Stark *et al.* 2015).

Finally, phosphatases regulate the G2/M transition and progression through mitosis. In *S. cerevisiae*, inhibition of PP2A function leads to cells arresting in G2, with a DNA content of 2N (Evans & Stark 1997). In *S. cerevisiae* PP2A dephosphorylates the Cdc25 phosphatase (Mih1 in *S. cerevisiae*) which promotes (although is not required for) entry into mitosis (Pal *et al.* 2008). It is worth noting that in *S. pombe* and higher eukaryotes, PP2A inhibits rather than promotes the G2/M transition (Kinoshita *et al.* 1993; Lee *et al.* 1991; Picard *et al.* 1989). In *Xenopus*, PP1 further stimulates Cdc25 (Margolis *et al.* 2003, 2006), and Cdc25 activity increases levels of M phase cyclins (Yano *et al.* 2013). In *S. cerevisiae*, Cdc14 phosphatase inhibits the G2/M transition in a manner akin to PP2A in *S. pombe* and higher eukaryotes (Mocciaro & Schiebel 2010). Inhibition of PP2A bound to its subunit Cdc55 is required for anaphase onset in *S. cerevisiae* (Lianga *et al.* 2013).

1.6 Replication origin characteristics

The intrinsic ability of a DNA sequence to recruit the initiator protein (ORC in eukaryotes) and subsequent pre-RC proteins determines whether it is able to function as an origin – this is the competence of an origin (section 1.2). Origins are licensed in G1 (section 1.4.1) and during S phase have the potential to be activated (section 1.4.2), according to a common molecular pathway. However, origins have other intrinsic characteristics which regulate activation, including the

point of origin activation within S phase (the replication timing of an origin) and how likely an origin is to fire in a given S phase (the efficiency of an origin). Origins do not act independently of one another, and clusters of origins which act to compensate for one another, or share similar characteristics have been described (Gillespie Blow 2010).

1.6.1 Origin timing

It was shown almost 60 years ago in mammals that some parts of the genome replicate early in S phase, and others late (Taylor 1960). Since then it has been confirmed that early and late replication domains exist in most eukaryotic genomes. Replication domains of 75 – 250 kb are found in *Drosophila* and *Arabidopsis* genomes, and domains span up to several Mb in mammalian cells (Hiratani *et al.* 2008; MacAlpine *et al.* 2004; Ryba *et al.* 2010). Early or late domains, referred to as constant timing regions (CTRs) are separated by timing transition regions (TTRs) (Desprat *et al.* 2009; Farkash-Amar *et al.* 2008; Hiratani *et al.* 2008; MacAlpine *et al.* 2004; Schübeler *et al.* 2002; White *et al.* 2004; Woodfine *et al.* 2004). The *S. cerevisiae* genome has not been found to have the same large-scale domains, however, its genome is smaller than the size of some mammalian replication domains. Despite this, broad zones of replication have been described as early or late replicating in yeast, and each replication origin has a characteristic activation time in S phase (Figure 1.11) (McCune *et al.* 2008; Raghuraman *et al.* 2001; Yabuki *et al.* 2002).

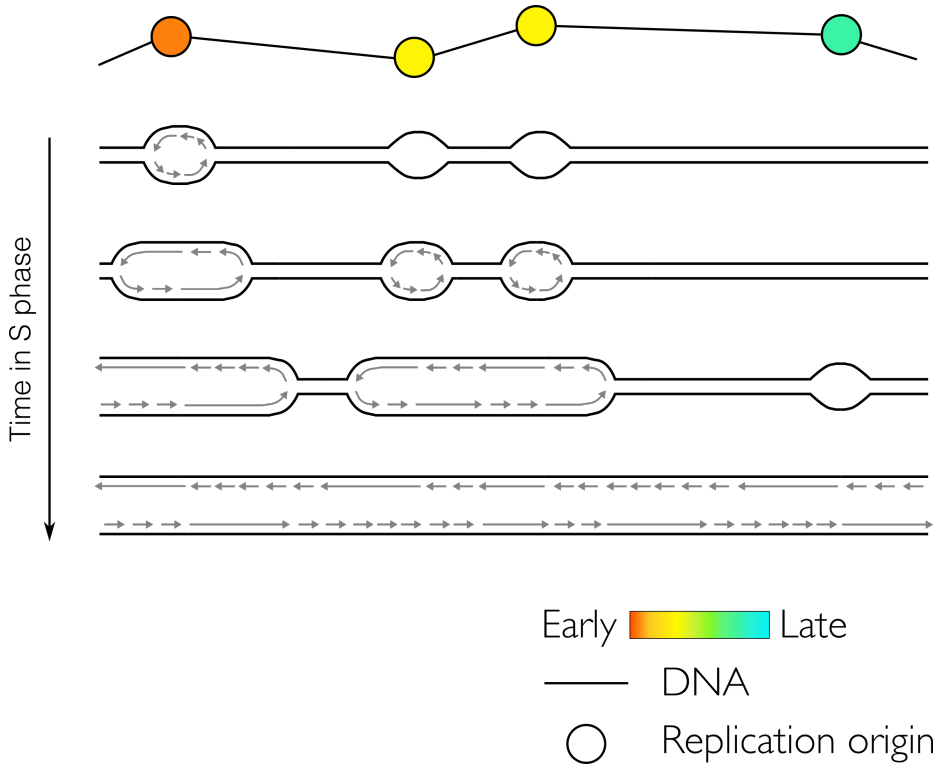


Figure 1.11 Schematic of DNA replication origin timing. *Eukaryotic DNA replication origins fire at characteristic, reproducible times during S phase. A representative genome area is shown (top) with origins represented by circles. Origin firing times are indicated by their fill colour (see key). DNA molecules over the genomic area are shown below, progressing through S phase from early to late. DNA surrounding early activating origins exists at a copy number of two for the majority of S phase, while DNA surrounding late activating origins exists at a copy number of one for the majority of S phase.*

The reproducible replication timing programmes which result from defined early and late replication domains are robust, and highly conserved in related eukaryotic species (Agier *et al.* 2018; Farkash-Amar & Simon 2010; Müller & Nieduszynski 2012; Rhind & Gilbert 2013; Ryba *et al.* 2010, 2011; Xu *et al.* 2012; Yang *et al.* 2018). Domains contain clusters of two to ten origins, and it has been

proposed that activation of one origin may prevent other origins in the domain from firing (Cayrou *et al.* 2011; Gillespie & Blow 2010). The dormant origins in a cluster are able to compensate for events such as fork stalling, preventing genome instability (Blow *et al.* 2011; Debatisse *et al.* 2012; Letessier *et al.* 2011).

Chromatin organisation contributes to replication timing in metazoans, with closely associated origins having similar replication activation times (Ryba *et al.* 2010). Replication domains have been shown to associate with chromatin loops (Buongiorno-Nardelli *et al.* 1982; Courbet *et al.* 2008) and topologically associated domains (TADs) (Dixon *et al.* 2012; Moindrot *et al.* 2012; Nora *et al.* 2012; Pope *et al.* 2014). Despite this, TADs can be uncoupled from replication timing (Ferreira & Carmo-Fonseca 1997; Ke *et al.* 2017). However, there is a broader division between early and late replicating DNA, where early replicating metazoan DNA is found in the centre of the nucleus and later replicating DNA is found at the nuclear periphery (Nakamura *et al.* 1986; O'Keefe *et al.* 1992; Solovei *et al.* 2009). Replication domains are also spatially defined within the nucleus in both metazoans and yeast (Lieberman-Aiden *et al.* 2009; Noble *et al.* 2011).

Epigenetic marks that define metazoan origins were discussed in section 1.2.4. These epigenetic marks (broadly associated with open chromatin) are enriched at early origins. For example, in yeast, an increase in histone acetylation, achieved by removing histone deacetylase complexes (HDACs) or tethering histone acetyltransferases (HATs) to late-activating origins results in the origins activating earlier in S phase (Aparicio *et al.* 2004; Knott *et al.* 2009; Vogelauer *et al.* 2002). Similar experiments carried out in human cells show this is a conserved feature of origin activation (Goren *et al.* 2008). The importance of

epigenetic marks is highlighted by observations that different alleles of some imprinted genes have different replication timings (Kawame *et al.* 1995; Kitsberg *et al.* 1993). Additionally, the DNA methylation state of rDNA repeats is important for their replication timing in human cells (Schlesinger *et al.* 2009). The local chromatin environment may be more important than any *cis*-acting factors, as transferring early origins to late-replicating domains results in those origins firing late in S phase (Friedman *et al.* 1996). The impact of epigenetic marks may depend on the recruitment of other factors. For example, the forkhead 1 and 2 (Fkh1 and Fkh2) transcription factors, required for the early-activation of some origins (Knott *et al.* 2012), bind via Cdc45 to early activing origins, and are proposed to be involved in Dbf4 recruitment (Fang *et al.* 2017).

Some regions of the genome have characteristic replication timing. Centromeres are early replicating in yeast, and in the yeast *Candida albicans* centromeres contain origins which are the first to activate in the genome (Koren *et al.* 2010). The early replication of centromeres contributes to the early replication of pericentric regions (Pohl *et al.* 2012). Association of kinetochore components may be required for early activation of replication at centromeres; some kinetochore mutants impact on the early replication time of centromeres (Batrakou *et al.* 2018; Natsume *et al.* 2013).

Conversely, telomeres are characteristically late-replicating. It was found that localisation of the telomeres to the nuclear periphery was neither sufficient nor required for their late-replication (Hiraga *et al.* 2006; Lian *et al.* 2011). However, a loss of proteins including Ku and Sir, which are required for telomere organisation leads to fewer telomere repeats and earlier replication timing (Cosgrove *et al.* 2002; Lian *et al.* 2011). Conversely, the Rap1 interacting factor

(Rif1) bound to the protein phosphatase 1 (PP1) limits telomere repeat length (Kedziora *et al.* 2018). However, the impact of Rif1 ablation on telomere repeat length and replication timing can be separated, indicating that this is not the only mechanism regulating telomere replication timing (Kedziora *et al.* 2018). The heterochromatic state of telomeres, resulting in closed chromatin which is hard for replication factors to access probably contributes to the late-replication of telomeres.

Some form of reproducible replication timing programme is seen in all eukaryotes. What advantage, if any, does this offer? There are potential advantages to limiting the number of ongoing replication forks at a given point in S phase. These include efficient use of the limiting concentration of dNTPs and the consolidation of a response to replicative stress to ensure the genome is fully replicated (Mantiero *et al.* 2011; Rivera-Mulia & Gilbert 2016). However, this does not explain why the timing programme is fixed. It has been shown that gene dosage of cell cycle genes, such as histones H2A and H2B, is increased in S phase due to their early replication (Müller & Nieduszynski 2017). However, the majority of genes are subject to a control pathway which prevents gene dosage effects (Padovan-Merhar *et al.* 2015; Voichek *et al.* 2016). Point mutations have been shown to occur at higher frequency in late-replicating regions of the genome, so there may be some advantage in replicating essential genes early (Supek & Lehner 2015). Early centromere replication in yeast is important for the stable segregation of chromosomes (Natsume *et al.* 2013). However, the question of why genomes replicate with fixed replication timing patterns remains to be fully explored.

1.6.2 Origin distribution and efficiency

Origins are on average spaced every 30 – 40 kb apart in the *S. cerevisiae* genome (Newlon *et al.* 1991; Shirahige *et al.* 1993; Siow *et al.* 2012). This is a higher frequency than is required to replicate the genome, so not every origin activates in each cell cycle (Friedman *et al.* 1997; Lebofsky *et al.* 2006; Yamashita *et al.* 1997). The percentage of S phases in which replication initiates from an origin is termed its efficiency. The efficiency of origins varies; in *S. cerevisiae* values range from less than 5 % to 90 % (Figure 1.12) (Friedman *et al.* 1997; Poloumienko *et al.* 2001; Yamashita *et al.* 1997).

One factor that can contribute to an origin being inefficient is activating late in S phase, and therefore being passively replicated by neighbouring origins prior to firing (Santocanale & Diffley 1996; Vujcic *et al.* 1999). However, these dormant origins are used when passive replication is prevented (Santocanale *et al.* 1999). A second factor that can impact on origin activation is origin interference. Although origin replication timing is set during G1 phase (Raghuraman *et al.* 1997), the activation of nearby origins can impact on other origins, even if they would typically fire at similar times in S phase (Brewer & Fangman 1993; Dubey *et al.* 1994).

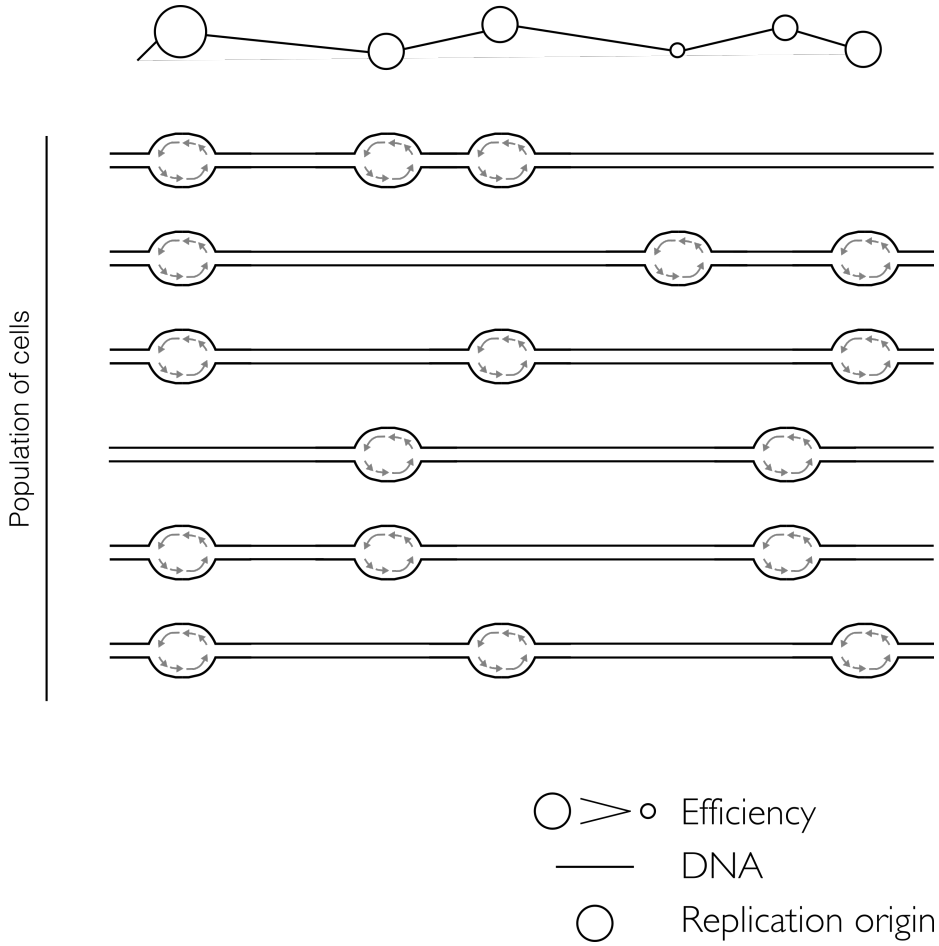


Figure 1.12 Schematic of DNA replication origin efficiency. *Eukaryotic DNA replication origins do not activate in every cell cycle. A representative genome area is shown (top) with origins represented by circles. Origin efficiencies are indicated by the circle size (see key). DNA molecules over the genomic area are shown below, from a population of cells. Efficient origins activate in most cells, while inefficient origins activate in a minority of cells.*

In metazoans, origins activate between 50 kb to 300 kb apart. Less than one in five origins fire in each cell cycle, which is logical given the two to ten origins which cluster in each replication initiation zone (Cayrou *et al.* 2011; Gillespie & Blow 2010). Each metazoan origin has an efficiency between 5 % and

20 %, with no origin as efficient as the most efficient *S. cerevisiae* origins (Blow & Ge 2008; Hamlin *et al.* 2008; Lebofsky *et al.* 2006). Metazoan origins activate in a more stochastic manner, which may allow for changes in origin efficiency depending on developmental stage or growth conditions (Méchali 2010). Origin interference is observed in human cells; when one origin is stochastically activated neighbouring origins are less likely to subsequently activate (Lebofsky *et al.* 2006).

1.7 Implications of de-regulated DNA replication

1.7.1 Over-replication

Re-activation of an origin in a single cell cycle leads to re-replication of the surrounding DNA, or over replication (Figure 1.13). To prevent re-replication, origin licensing is limited to G1, and origin firing is restricted to S phase (Green *et al.* 2010). In both yeast and mammalian cells, inhibitors function to prevent replication initiation factors acting aberrantly (Hook *et al.* 2007; Suryadinata *et al.* 2010). However, origin licensing proteins are inappropriately expressed in the early stages of many cancers (Blow & Gillespie 2008; Hiraiwa *et al.* 1997). Testing for the overexpression of origin licensing factors MCM and Cdc6 can reduce false-negative rates in cervical smears (Williams *et al.* 1998).

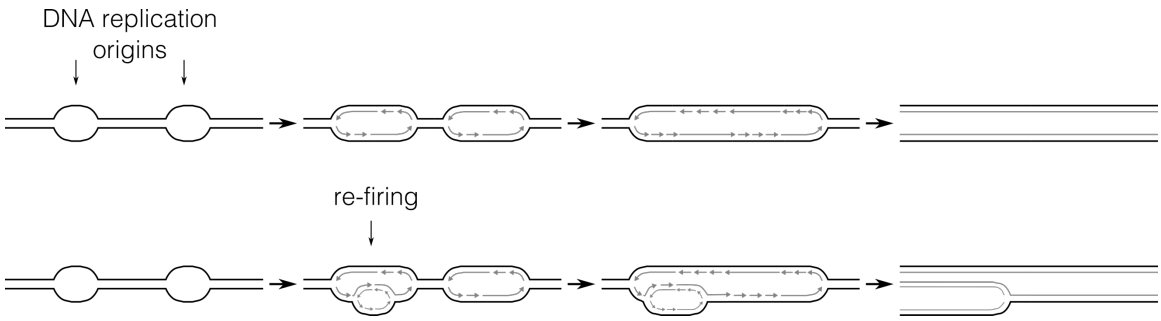


Figure 1.13 Schematic of over-replication. *Overexpression of origin licensing proteins or a lack of origin regulation can result in re-activation of an origin in the same cell cycle. Origin re-activation leads to re-replication of the surrounding DNA. Over-replicated DNA is a hallmark of genome instability.*

Overexpression of Cdt1 leads to re-replication, producing cells with higher DNA content and erroneous dsDNA fragments (Davidson *et al.* 2006; Nishitani *et al.* 2004). A p53-dependent DNA damage pathway acts to prevent Cdt1-dependent re-replication in metazoans (Vaziri *et al.* 2003). However, Cdt1 has also been shown to cause genome instability without re-replication (Tatsumi *et al.* 2006). Conversely, inhibition of the early mitotic inhibitor 1 (Emi1) protein leads to ssDNA and re-replication which converts ssDNA to double strand breaks (DSBs) (Neelsen *et al.* 2013).

Loss of DNA re-replication control leads to many forms of genome instability (Alexander & Orr-Weaver 2016; Orr *et al.* 2010). The genome damage observed is consistent with head-to-tail collisions of adjacent replication forks (Davidson *et al.* 2006). Additionally, re-replication forks demonstrate reduced elongation, traversing only 30 – 35 kb from the origin (Nguyen *et al.* 2001). These re-replication forks are prone to breakage which can result in gene amplification (Finn & Li 2013; Green *et al.* 2010).

1.7.2 Under-replication

In the absence of origin firing, DNA under-replication can occur (Figure 1.14). Genomic instability is induced when levels of replication factors, such as MCM, are insufficient in proliferating cells (Elble & Tye 1992). MCM mutations confer elevated levels of chromosomal rearrangement (Dershowitz *et al.* 2007; Hawkins *et al.* 2013a). Despite the lethal effects anticipated when DNA is unable to replicate, origin deletion has been found to be tolerated in a range of model organisms (Hyrien *et al.* 2003). This tolerance is due, at least in part, to the abundance of origins able to compensate for replication initiation (section 1.6.2).

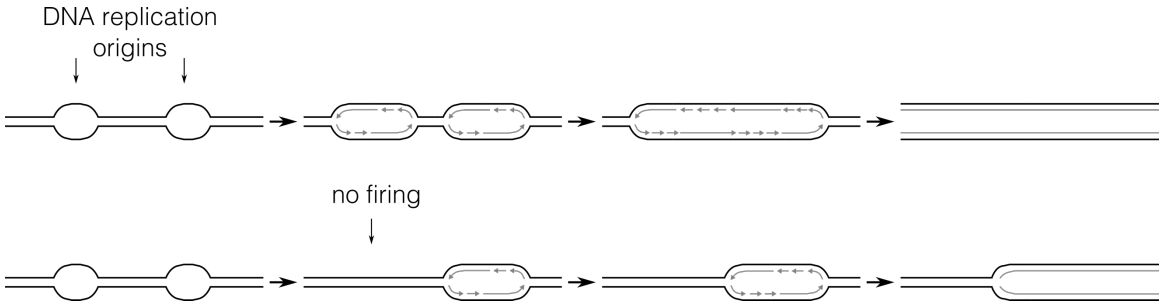


Figure 1.14 Schematic of under-replication. *A lack of origin activation or mutations of origin activation factors can result in regions of the genome without origin activation. Insufficient origin activation results in under-replication of the surrounding DNA. Under-replicated DNA is a hallmark of genome instability.*

However, a paucity of origin activation events, in a chromosomal region or genome wide, leads to replicative stress and increases the risk of chromosome breakage (Blow *et al.* 2011; Letessier *et al.* 2011). Human syndromes associated with insufficient origin licensing (thereby preventing origin activation) include

some cancers, adrenal insufficiency and natural killer cell deficiency resulting from MCM mutations, dwarfism phenotypes and Meier-Gorlin syndrome from ORC mutations and Fragile-X syndrome resulting from a scarcity of origins over the fragile site (Bicknell *et al.* 2011; Gineau *et al.* 2012; Guernsey *et al.* 2011; Hughes *et al.* 2012a; Kunnev *et al.* 2010; Letessier *et al.* 2011; Pruitt *et al.* 2007; Shima *et al.* 2007). When origin licensing is reduced, dormant origins become key for maintaining genome stability (Kawabata *et al.* 2011; Klotz-Noack & Blow 2011).

The high level of regulation which cells dedicate to controlling DNA replication origin licensing and the serious consequences of depleted replication licensing suggest genomes may employ alternative replication mechanisms to avoid under-replication. These alternative methods of replication are discussed in section 1.8.

1.8 Alternative methods of DNA replication

Despite an overabundance of origin sites to facilitate complete replication, the stochastic nature of origin firing means that large replicons that are unreplicated are predicted to occur frequently: this is the random gap problem (Kreuzer 2000). Alternative replication initiation pathways can help circumvent this problem. Currently known alternative replication mechanisms can be broadly divided into two categories: those where non-replicative processes are co-opted to initiate replication from aberrant DNA structures, and those where the canonical pathway functions in non-canonical settings.

1.8.1 Bacteria

1.8.1.1 Induced stable DNA replication

Break-induced replication was first described in the T4 bacteriophage; it promotes replication at the ends of linear chromosomes by invading homologous sequences present on other chromosome termini, forming a D-loop at which a replication fork assembles (Asai *et al.* 1993). Initially thought to be anomalous, D-loop-dependent replication is now known to be essential for *E. coli* genome replication. When the SOS response is constitutively activated, *oriC* replication is inhibited, but *E. coli* can still replicate their chromosome, dependent on RecA recombinase; this was coined induced stable DNA replication (iSDR) (Asai *et al.* 1994; Kuzminov & Stahl 1999). In iSDR, strand invasion at D-loops, mediated by the RecBCD helicase (Burland *et al.* 1993), occurs at double-strand ends which are protected by chi -sequences. Chi-sequences are over-represented in the *E. coli* genome, orientated away from *oriC* (Liu *et al.* 1999); this encourages recombination-dependent replication to initiate co-transcriptionally. The PriA primosome protein facilitates DnaB loading at D-loops (Liu *et al.* 1999). Strains with null *priA* have a massively reduced viability and rapidly gain suppressor mutations in DnaC, the DnaB loader. Mutations allow PriA-independent loading of DnaB to D-loops (Kogoma 1978). Collectively, this evidence suggests that D-loop dependent replication is an essential replication mechanism in *E. coli*.

1.8.1.2 Constitutive stable DNA replication

Canonical DNA replication requires *de novo* replication factor synthesis, however stable DNA replication (*sdr*) mutants demonstrate constitutive replication (de Massy *et al.* 1984; von Meyenburg *et al.* 1987). Constitutive stable DNA

replication (cSDR) is independent of *oriC*, DnaA and cell cycle stage. cSDR alone can enable cellular growth, although this growth is at a reduced rate (Ogawa *et al.* 1984). However, cSDR initiation depends on the absence of RNaseH (*rnh*), which processes RNA primers (Dasgupta *et al.* 1987; Masukata *et al.* 1987). In *rnh* cells unprocessed RNA can hybridise to DNA, exposing a 40bp ssDNA region: this R-loop initiates replication (de Massy *et al.* 1984; Maduike *et al.* 2014). Kogoma and colleagues proposed cSDR initiated at discrete loci – *oriK*s - although recent mapping by next generation sequencing suggests initiation site flexibility (von Meyenburg *et al.* 1987). Like iSDR, cSDR seems to occur simultaneously with canonical DNA replication under normal cellular conditions (Hong *et al.* 1995).

Mutants in RecG recombinase also facilitate cSDR (Rudolph *et al.* 2013). However, initiation occurs only in the terminator region of *E. coli dnaA recG* mutants. Due to uni-directional replication fork traps surrounding the terminator region (*ter*/Tus replication fork trap) *dnaA recG* mutants are unviable unless these trap sequences are also removed (Dimude *et al.* 2015). Initiation is restricted to the termination region because, unlike R-loop-dependent replication, *recG*-dependent replication is initiated by free 3' flaps generated by replication fork fusion (Dimude *et al.* 2015). These 3' flaps are processed by recombination proteins including PriA, suggesting an initiation mechanism more analogous to iSDR than that seen in *rnh* mutants (Dimude *et al.* 2015).

1.8.1.3 A less favoured approach?

Although both iSDR and cSDR are capable of maintaining cell viability, DNA replication origins have been retained during bacterial evolution. A single origin

which replicates bi-directionally to the terminus region allows transcription in *E.coli* to co-align with replication. Conversely, bi-directional replication from other loci leads to transcription and replication complex interactions. Collisions require resolution by recombination proteins and increase the risk of genomic instability (Usongo *et al.* 2016). Mutating *rnh* strains to further increase R-loop formation leads to DNA relaxation and growth defects, despite functional canonical initiation from *oriC*. Defects are resolved by suppressor mutations in cSDR-specific replication machinery (Ausiannikava & Allers 2017). This indicates that, whilst alternative replication mechanisms can function alongside canonical processes in bacteria, there is a selective advantage in replication initiating at *oriC*.

1.8.2 Archaea

1.8.2.1 Initiator dependence

Since many archaeal species include more than one origin per chromosome, multiple initiator proteins are common (Coker *et al.* 2009; Samson *et al.* 2013). In both *Sulfolobus islandicus* and *Halobacterium* species NRC-1 it has not been possible to inactivate all Orc1/Cdc6 copies in the genome suggesting viability in these species depends on canonical replication initiation (Samson *et al.* 2013). However, the effects of deleting initiator proteins cannot be equated to the effects of deleting replicator sequences. In *S. islandicus*, only two of three Orc1/Cdc6 homologues associate with origins; these can be co-deleted (Samson *et al.* 2013). Whilst it was not possible to additionally delete the third Orc1/Cdc6 gene, this gives stronger evidence for a general Orc1/Cdc6 requirement than an origin-binding requirement (Yang *et al.* 2015).

1.8.2.2 Dormant origin activation

Conversely, when active origin sequences were inactivated in *Haloferax mediterranei*, replication became dependent on a fourth, dormant origin (Wu *et al.* 2012b; Yang *et al.* 2015). Dormant origin sequences have also been discovered in *Haloferax volcanii* and *Haloarcula hispanica* (Yang *et al.* 2015). An inability to remove all four *H. mediterranei* origins simultaneously provides strong evidence for replicator dependence in this species (Hawkins *et al.* 2013a).

1.8.2.3 Recombination based initiation

In contrast, *H. volcanii* demonstrates a remarkable capacity to survive despite inactivation of all its main chromosomal origins, contrary to expectation, growth without origins is enhanced (Hawkins *et al.* 2013a). Replication timing data show no dormant origins are activated. Instead, survival becomes dependent on RadA, the archaeal RecA homologue (Norais *et al.* 2007). The capacity for recombination-dependent replication may be due in part to the high and variable ploidy of *H. volcanii*, and also to its containing a plasmid, pHV2, which replicates in a recombination-dependent manner (Hawkins *et al.* 2013a). Although bearing resemblance to bacterial iSDR replication, no facilitating mutations (such as those activating an SOS-like response) are required for origin-independent replication of *H. volcanii* (Ausiannikava & Allers 2017).

RecA in bacteria also protects the replication fork by promoting fork reversal into a “chicken-foot” structure, which can lead to branch migration and DNA replication with the aid of RuvABC proteins (Gowrishankar 2015). However, since this mechanism requires an ongoing replication fork, it is more likely that

the RadA required for origin-independent replication of *H. volcanii* functions to promote recombination.

Dependence on RadA indicates that no RadA-independent alternative replication can maintain viability in *H. volcanii* (Hawkins, 2013). Further, no evidence for archaeal replication initiated from RNA:DNA hybrids has yet been presented. Due to the diversity of replication machinery in archaea, studies in multiple models are required to elucidate all alternative replication mechanisms (Nieduszynski *et al.* 2006). However, as indicated here, studies in multiple archaeal species have shown dependence on the canonical initiation pathway, or adaptions thereof.

1.8.3 Budding yeast

S. cerevisiae utilises alternative replication initiation methods seen both in bacteria and in archaea: activation of dormant origins and non-canonical initiation (Dershowitz *et al.* 2007). In budding yeast origins occur every 30 - 40 kb in the genome, however only small subsets of these are activated in each cell cycle. The apparent excess of origins may permit typically inactive, 'dormant', origins to rescue under-replicated DNA. Non-canonical initiation, defined as non-ACS-dependent, appears to still utilise canonical initiation factors: it is ORC-dependent (On *et al.* 2014). Origins are not required for DNA replication to initiate: *in vitro*, ORC and pre-replication complex (RC) components can replicate plasmid DNA without ACS sites, so long as excess kinase and firing factors are present. However, origin-dependency is seen under physiological conditions when competitor DNA is present, or ORC concentrations are limited (Morrow *et al.* 1997). This is reminiscent of the situation *in vivo*.

1.8.3.1 Break-induced replication (BIR)

Successful transformation of yeast with telomere-deficient linear constructs first demonstrated that recombination events could lead to sequence replication (Dunn, 1984). A 'break-copy' mechanism was proposed by Morrow in 1997, where a uni-directional replication fork proceeded from a DNA break site towards telomeres (Ira & Haber 2002). In this process, now known as break-induced replication (BIR), Rad51 binds ssDNA, exposed by resection of DSBs, and invades DNA with a large (~100 bp) region of homology (Donnianni & Symington 2013; Saini *et al.* 2013). This creates a D-loop which can migrate to the end of a chromosome (Figure 1.15) (Ira & Haber 2002). Rad51-independent BIR can also occur; this utilises an intramolecular single-strand annealing mechanism, requiring smaller (30 bp) regions of homology. However, Rad51-independent mechanisms account for only 10 % of BIR (Wilson *et al.* 2013).

In Rad51-dependent BIR the Pif1 helicase is essential, alongside Pol32 (a Polδ subunit). Pif1 liberates the nascent strand, preventing excessive torsion, facilitating the migration of the D-loop, essential for long replication tracts (Donnianni & Symington 2013). Despite the capacity for over 100 kb of DNA to be synthesised, BIR is not efficient if the site of initiation is greater than 20 kb from a telomere (Saini *et al.* 2013). Additionally, lagging strand synthesis is delayed relative to the leading strand (Deem *et al.* 2011; Smith *et al.* 2007), which leaves ssDNA exposed (Figure 1.15). Therefore, BIR is highly mutagenic, both proximal and distal to the initiation site; frameshift mutations are increased up to 2800 times, and template switching due to unstable D-loops leads to gene conversions and chromosomal rearrangements (Mayle *et al.* 2015).

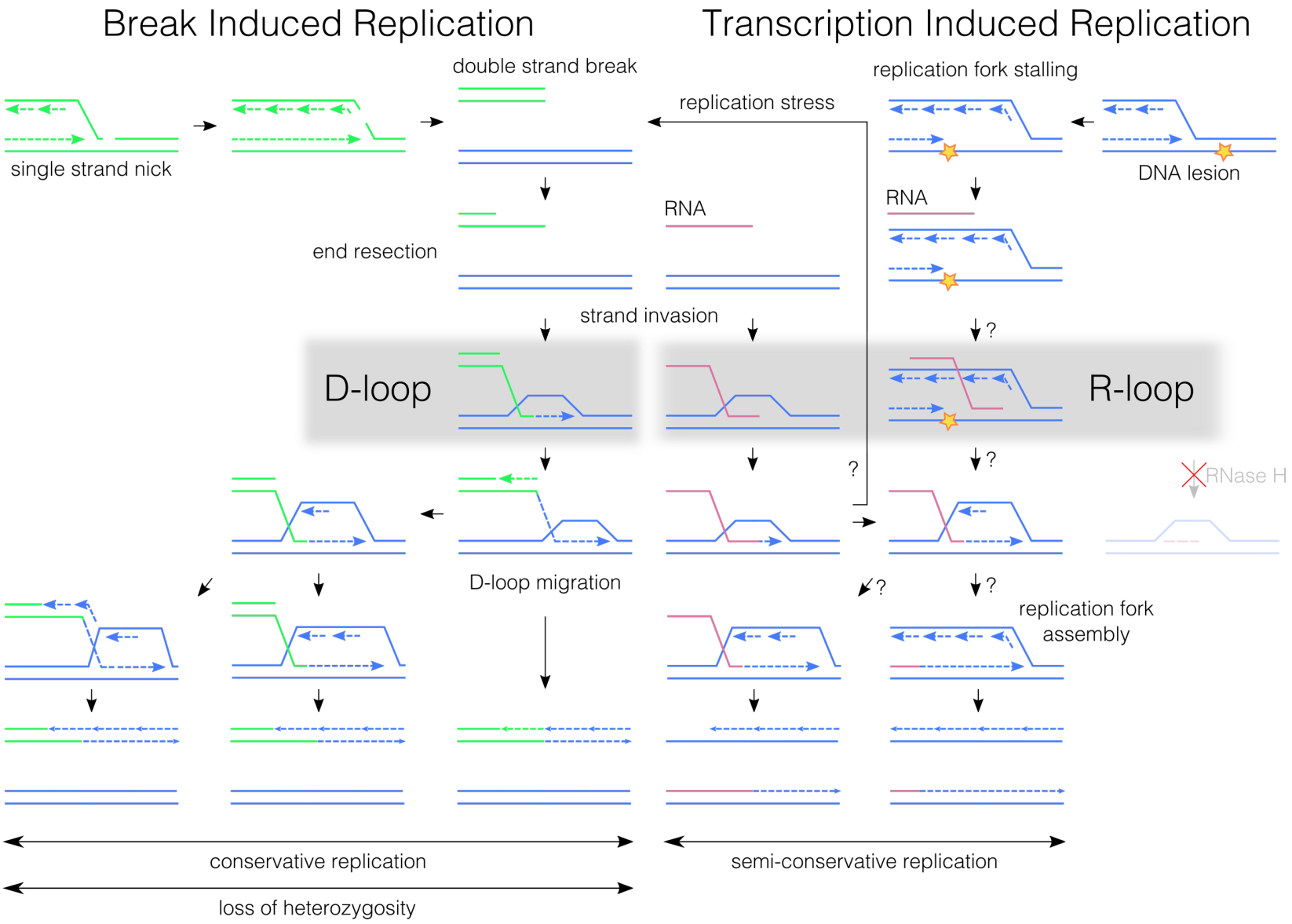

Break Induced Replication
Transcription Induced Replication
double strand break
replication fork stalling
replication stress
single strand nick
DNA lesion
RNA
RNA
end resection
strand invasion
?
D-loop
R-loop
?
?
RNase H
D-loop migration
?
?
replication fork
assembly
conservative replication
semi-conservative replication
loss of heterozygosity

Figure 1.15 A summary of Break-induced Replication (BIR) and Transcription-induced Replication (TIR). *The DNA duplex (blue, solid) can be invaded by free 3' DNA ends (green, solid) produced from resection at a double strand break (DSB) to create a D-loop. Alternatively, strand invasion by RNA (pink, solid) can occur, forming an R-loop. Post strand-invasion, new DNA is synthesised (dashed). D-loops migrate, with delayed lagging stand synthesis. the mechanism for R-loop mediated synthesis has not been established: conversion to a DSB, migrating R-loops and* de novo *assembly or restart of a stalled replication fork are possible.*

Recently it was shown that Mus81, capable of converting a D-loop to a replication fork by incision (Ehmsen & Heyer 2008), can supress the mutagenic effects of BIR which initiates at broken replication forks (Mayle *et al.* 2015). Mus81 activity restricts increased mutation rates to within 10 kb of the break site and prevents template switching (Mayle *et al.* 2015). Oncoming replication forks also supress BIR-dependent mutagenesis, presumably by displacing BIR machinery (Chan *et al.* 2014; El Hage *et al.* 2014). Therefore, whilst BIR can occur simultaneously with canonical replication initiation, and rescue recombination intermediates from DSBs, the cell preferentially converts D-loops to a full replication fork via Mus81, or replicates the DNA canonically

1.8.3.2 Transcription-induced replication (TIR)

In prokaryotes iSDR and cSDR were discovered in tandem, however, in eukaryotes evidence for transcription-induced replication (TIR) has been less forthcoming. R-loops were revealed in the budding yeast genome only in 2014,

primarily in ribosomal DNA, although when RNaseH enzymes are inhibited also in tRNA, retrotransposons and even actively transcribed protein coding genes (Stuckey *et al.* 2015). RNaseH1 functions to remove DNA:RNA hybrids in mitochondrial DNA and is essential for mitochondrial DNA replication (Cerritelli *et al.* 2003), whereas RNaseH2 functions predominantly in the nucleus to remove DNA:RNA hybrids (Frank *et al.* 1998). In the absence of RNaseH1/2 and TopI, R-loops within 35S rDNA can initiate transcription, although this occurs after canonical replication (Arora *et al.* 2014). It is also speculated that telomeric repeat-containing RNA (TERRA), a non-coding telomere-associated RNA, may lead to TIR; TERRA is capable of forming RNA:DNA hybrids that are regulated by RNase H1 (Arora *et al.* 2014). However, RNase H1 depletion decreases replication initiation in telomeres and causes R-loop-mediated replication stress: strong evidence to the contrary (Rippe & Luke 2015). Nevertheless, replication can initiate in telomeres and TERRA may aid in replication restart, either by recruiting recombination factors or re-priming a replication fork in a manner akin to cSDR (Aguilera & Garcia-Muse 2012; Ravoitytė & Wellinger 2017).

The mechanism for TIR in eukaryotes is not known. The involvement of TERRA hints at an R-loop conversion to DSB mechanism, which may recruit BIR factors for repair (Figure 1.15). Since there is an overlap between many iSDR and cSDR essential factors, including the primosomes and RecA, it is feasible that TIR and BIR are interlinked in this fashion: the Pif1 helicase, essential for D-loop migration, can also unwind R-loops (Ravoitytė & Wellinger 2017). However, other mechanisms including direct replication fork restart after R-loop mediated collapse, or even *de novo* assembly of a replication fork are possible: TIR does

not lead to conservative replication (Figure 1.15) (Aguilera & Garcia-Muse 2012; Stuckey *et al.* 2015).

R-loop driven replication is dangerous for cells, as RNA:DNA hybrids are mutagenic (Stuckey *et al.* 2015). Therefore, like cSDR, the *in vivo* use of this pathway is minimised, and in order to study TIR, a mutant *rnh1*Δ*, rnh2*Δ background is required (Stuckey *et al.* 2015). Nevertheless, RNA:DNA hybrids are seen at low levels in wild-type strains (Lombraña *et al.* 2016). Particularly, areas of high transcription can prompt replication fork stalling; to avoid this the replication timing of the *Leshmania* parasite is co-ordinated with its transcriptional profile (Dershowitz *et al.* 2007). TIR-restart in highly transcribed areas may be essential to ensure genome replication is completed.

1.8.3.3 Origin-less replication in yeast

Prior to the discovery of TIR in eukaryotes, the capacity for yeast to maintain linear chromosome fragments lacking active origins was established (Theis *et al.* 2007). An initial investigation of 'origin-less' fragment maintenance (OFM) uncovered two mechanisms for replication: activation of dormant origins, and an alternative form of replication, which was recombination-independent (Theis *et al.* 2010). The first gene identified as necessary for OFM was *RAD9* (*ofm*14), which intersects the DNA damage and replication stress response pathways. A larger genetic screen unearthed more than 50 *ofm* genes. These clustered into DNA damage response and DNA damage signalling pathways, indicating that the replication stress pathway was dispensable (Theis *et al.* 2007). However, a separation of function study indicated the majority of the *ofm* phenotype of *rad9* was due to loss of dormant origin activation, not alternative replication (Theis *et*

al. 2010). Further, Rad9 is now known to be involved in a cell cycle checkpoint that causes the cell to wait for replication to finish, further supporting the notion that Rad9 has no role in an alternative pathway for replication (Ivanova *et al.* 2018).

In the light of current knowledge, a re-examination of *ofm* genes can illuminate potential alternative replication mechanisms (Fasching *et al.* 2015; Theis *et al.* 2010). For example, *POL32* is required for OFM, suggesting a BIR-like mechanism. However, lack of Rad52-dependence, and the additional requirement for *TOP3*, which dissolves D-loops, preventing their migration, precludes a recombination-driven mechanism (Theis *et al.* 2010). Conversely, requirement of *RPA34*, a subunit for RNA polymerase I, indicates that efficient transcription is required (Stuckey *et al.* 2015). This lends itself to a TIR-like mechanism: replication initiation at RNA:DNA hybrids is RNA polymerase I dependent (Dershowitz *et al.* 2007). Further, a requirement for linear fragments suggests that telomeres assist in this process (Theis *et al.* 2010). Additional *ofm* proteins Swi5, a transcriptional activator, and Whi4, an RNA binding protein, emphasise the likely role of transcription in OFM (Dershowitz *et al.* 2007).

1.8.3.4 Non-specific canonical initiation

In the absence of active origins, dormant origins, which are less efficient or later-activating, can facilitate replication (Bogenschutz *et al.* 2014). Moreover, replication of yeast chromosome VI, despite inactivation of origins on its left arm, demonstrates replication initiation from non-canonical sites (Gros *et al.* 2014; On *et al.* 2014). Additional, *in vitro*, studies reveal that ORC has the capacity to bind both ACS variants and non-specific sequences in the absence of competitor DNA

(Devbhandari *et al.* 2017; Kurat *et al.* 2017). Nevertheless, in the context of chromatin ORC is restricted to ACS sites (Gros *et al.* 2015; Remus *et al.* 2009). However, post-loading, MCM helicases are able to slide along DNA and initiate replication from non-ACS sites (Eaton *et al.* 2010). Therefore, whilst essential for spatial organisation of DNA replication, origins do not appear to be intrinsically required for DNA unwinding.

1.8.4 Other eukaryotes

In yeast, the ACS can exclude nucleosomes, facilitating ORC binding to DNA (Eaton *et al.* 2011). This highlights the importance of non-sequence origin determinants. In multicellular eukaryotes, where no conserved origin sequence has been defined, epigenetic marks, DNA secondary structure (e.g. G4 quadruplexes) and nucleosome positioning may define replication origins; epigenetic definition facilitates tissue-specific replication patterns (Nguyen *et al.* 2015).

Despite a less well-defined canonical process, evidence for non-canonical replication is found in other eukaryotes. In the fission yeast *S. pombe* homologous-recombination-restarted-replication (HoRReR) occurs. HoRReR is a mutagenic process due to the invading 5' DNA end continually dissociating from the DNA it invades (Miyabe *et al.* 2015). Unlike BIR in *S. cerevisiae*, HoRReR leads to semi-conservative replication, however, as in BIR, Polδ is required: it completes both leading and lagging strand synthesis (Cullen *et al.* 2007; Pietrobon *et al.* 2014). Further, Rad51 promotes template switching from homologous recombination intermediates, including D-loops. Combined with

break-induced loss of heterozygosity, this indicates that homology-driven strand invasion post-DSBs occurs in fission yeast (Lee & Clayton 1998).

The capacity in eukaryotic cells for transcription-induced replication is evidenced by mitochondrial DNA replication (McDonald *et al.* 2016). However, whether this aptitude is used to replicate genomic DNA is unclear. Phf1 helicase, the *S. pombe* Pif1 homolog, is required for replication of tRNA, highly transcribed genes and 5S ribosomal genes (Gagou *et al.* 2014). Moreover, disruption of human Pif1 negatively impacts replication fork restart (Lombraña *et al.* 2015). R-loops may act in CpG island promoters, exposing single-stranded replicator signals, such as G4 quadruplexes, to enable efficient replication initiation from origins (Nieduszynski *et al.* 2006).

1.9 Objectives of this study

This purpose of this thesis was to further the current understanding of DNA replication origin regulation, and how cells respond when their DNA replication timing is perturbed. This work was carried out in *S. cerevisiae*, an ideal model in which to study the regulation of origins due to the ACS motif being well characterised (Nieduszynski *et al.* 2007; Siow *et al.* 2012), and well annotated (e.g. in the OriDB database) (Yeeles *et al.* 2015). Additionally, there is high conservation between DNA replication initiation proteins in *S. cerevisiae* and multicellular eukaryotes, such as humans.

First, this thesis set out to examine control of origin firing. Until a few years ago, the replication initiation pathway, which begins with ORC binding to the origin ACS site, was considered to be a stepwise activation process (Yeeles *et al.*, 2015). DDK then subsequent CDK phosphorylation of MCM helicase and firing

factors respectively was though to allow assembly of the active CMG helicase on DNA and origin firing process (for review see: (Davé *et al.* 2014; Foti *et al.* 2016; Mattarocci *et al.* 2014)). However, recent work has shown phosphatases such as Rif1-PP1 can also play a key role in regulating DNA replication initiation (Davé *et al.* 2014; Foti *et al.* 2016; Mattarocci *et al.* 2014). A possible inhibitory role for the phosphatase subunit Rts1 during DNA replication initiation was investigated by looking for synthetic rescue of temperature-sensitive (ts-) replication initiation factors by Δ*rts1*.

Subsequently, this thesis aimed to examine the consequences for cell growth and DNA replication of de-regulating a series of origins in the *S. cerevisiae* genome. A *S. cerevisiae* strain with an unnaturally large replicon of over 100 kb was generated by inactivating ACS sites on the right arm of chromosome VI (ChrVI). Cell growth was measured by examining mean generation time of yeast strains. DNA replication was measured using high-resolution genomic methods which quantitatively examined relative DNA copy number over the large replicon.

Finally, this thesis aimed to examine the progression of replication forks across a large replicon of DNA. A next-generation sequencing approach, Polymerase-usage sequencing (Pu-seq), was utilised to examine replication fork direction. Pu-seq provided a quantitative method to examine leading and lagging strand polymerase usage, a proxy for replication fork direction, due to its strand-specific nature.

Ultimately, the objective of this thesis was to investigate the regulation of DNA replication origins and the consequences of de-regulated DNA replication origins for DNA replication.

2 Methods

2.1 Chemicals and enzymes

Chemicals routinely used in the laboratory were purchased from the following vendors: Thermo Fisher Scientific, Sigma, ChemBio, VWR, Melford Laboratories Ltd and Scientific Laboratory Solutions. All restriction enzymes and associated buffers were purchased from New England Biolabs (NEB). Other enzymes were purchased as following: *Taq* DNA polymerase and associated buffers from Bioline, NEB Gibson Assembly Cloning kit (E5510) from NEB and Zymolyase 20T and 100T from the Seikaguku Corporation.

2.2 Growth media

All media have their supplier and catalogue numbers listed. Media were made up according to the manufacturer's recommendations, autoclaved at 117 °C for 15 minutes, and then stored at room temperature unless otherwise stated.

Yeast-Peptone-Adenine-Dextrose (YPAD) medium

Formedium, CCM1010 (10 g/l yeast extract, 20 g/l peptone, 20 g/l glucose, 0.04 g/l adenine sulphate).

Yeast-Peptone-Adenine-Dextrose (YPAD) agar

Formedium, CCM0510 (10 g/l yeast extract, 20 g/l peptone, 20 g/l glucose, 0.04 g/l adenine sulphate, 20 g/l agar).

G418 containing medium for yeast (400 µg/ml)

G418 stock solution (100 mg/ml) was made from G418 disulfate salt (Sigma A1720-5G) and sterilized using a 0.45 µm filter. Cool (60 °C), autoclaved, YPAD agar or broth was supplemented with 4 ml/l of the G418 stock solution, giving a final concentration of 400 µg/ml.

Ampicillin containing medium for yeast (100 µg/ml)

Ampicillin stock solution (100 mg/ml) was made from ampicillin sodium salt (Sigma 9518-25g) and sterilized using a 0.45 µm filter. Cool (60 °C), autoclaved, YPAD agar or broth was supplemented with 1 ml/l of ampicillin stock solution, giving a final concentration of 100 µg/ml.

Sporulation medium (agar)

Formedium, CCM0802 (10 g/l yeast extract, 0.2 g/l glucose, 10 g/l potassium acetate, 20 g/l agar).

Minimal medium for yeast (agar)

Formedium, CSM0105 (contains a yeast nitrogen base without glucose or amino acids).

Minimal medium for yeast (broth)

Formedium, CSM0205 (contains a yeast nitrogen base without glucose or amino acids).

Complete medium for yeast (agar and broth)

Formedium, DCS0031 (composition listed below (Table 2.1)). Added to minimal agar or minimal broth (see above) to make complete medium.

Table 2.1. The composition of complete medium

AMINO ACID (ABBREVIATION)	*CONCENTRATION (mg/l)*
Adenine (ADE)	10
L-Arginine (ARG)	50
L-Aspartic acid (ASP)	80
L-Histidine HCl (HIS)	20
L-Isoleucine (ILE)	50
L-Leucine (LEU)	100
L-Lysine HCl (LYS)	50
L-Methionine (MET)	20
L-Phenylalanine (PHE)	50
L-Threonine (THR)	100
L-Tryptophan (TRP)	50
L-Tyrosine (TYR)	50
Uracil (URA)	20
Valine (VAL)	140

Synthetic dropout medium for yeast (agar and broth)

One of the supplement mixes listed in Table 2.2 was added to minimal agar or minimal broth (see above) to make complete media lacking one amino acid. The composition of complete medium is shown in Table 2.1.

Table 2.2. Supplements used to create synthetic dropout medium for yeast

AMINO ACID DROPPED OUT OF SUPPLEMENT (ABBREVIATION)	*CONCENTRATION (mg/l)*	*FORMEDIUM CATALOGUE NUMBER*
Leucine (-LEU)	700	DCS0231
Uracil (-URA)	800	DCS0281
Histidine (-HIS)	800	DCS0201
Lysine (-LYS)	770	DCS0241
Tryptophan (-TRP)	750	DCS0261

Luria-Bertani (LB) plates for *E. coli*

Sigma, L3147-250G (15 g/L agar, 10 g/L tryptone, 10 g/L sodium chloride, 5 g/L yeast extract).

Luria-Bertani (LB) broth for *E. coli*

Melford, L24040-1000.0 (10 g/l casein digest peptone, 10 g/l sodium chloride, 5 g/l yeast extract).

Ampicillin containing medium for *E. coli* (100 µg/ml)

Ampicillin stock solution (100 mg/ml) was made from ampicillin sodium salt (Sigma 9518-25g) and sterilized using a 0.45 µm filter. Cool (60 °C), autoclaved, LB agar or broth was supplemented with 1 ml/l of ampicillin stock solution, giving a final concentration of 100 µg/ml.

2.3 Yeast strains, plasmids and oligonucleotides

Most yeast strains used had a W303 background (Schacherer *et al.* 2007). However, strains from the *S. cerevisiae* genome deletion project (SGDP) (Giaever *et al*. 2002) had an S288c background. Additionally, some strains sourced from other laboratories had the backgrounds Δ|(-2)|-7B-YUNI300 or CN5-19a/CF4-16B33. All strains used are listed in Table 2.3.

Table 2.3. Yeast strains used in this study.

STRAIN	*BACKGROUND GENOTYPE*	*SOURCE*
T7107	W303 *MAT**a** RAD5 BUD4 leu2 ura3 trp1 ade2 his3*	T. Tanaka lab
K2539	*MATα cdc9-1*	T. Tanaka lab
dbf4-1	*MAT**a** ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 can1-100 ssd1-d2 Gal+ dbf4-1*	Tanaka & Nasmyth 1998
YKB2	*MAT**a** leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3- 11,15, cdc7-4*	Mattarocci *et al*. 2014
YYK32	*MAT**a** leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3- 11,15 cdc45-27 bar1Δ::hisG*	Mattarocci *et al*. 2014
YYK14	*MAT**a** leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3- 11,15 sld3-4 bar1Δ::hisG*	Mattarocci *et al*. 2014
YNIG63(2)	*MAT**a** leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3- 11,15 dpb11-24 bar1Δ::hisG*	Mattarocci *et al*. 2014
YCH175	W303 *MATα ho ade2 trp1 can1 leu2 his3 GAL psi + cdc6-1*	Mattarocci *et al*. 2014
YOR014W	S288c *MAT**a** rts1Δ::kanMX*	Giaever *et al*. 2002

YBR275C	S288c *MAT**a** rif1Δ::kanMX*	Giaever *et al*. 2002
YDR007W	S288c *MAT**a** trp1Δ::kanMX*	Giaever *et al*. 2002
CB08D3	CN5-19a/CF4-16B33 *MAT**a** trp1- 63 ura3-52 ade2-101 / balancer: his4-280 C2G::SUP11-1 H9G::TRP1 his4-S8 MAT**a***	Dershowitz *et al*. 2007
YDN289	CN5-19a/CF4-16B33 *MAT**a** rp1- 63 ura3-52 ade2-101 his4-280 C2G::SUP11-1 6ORI H9G::TRP1 /balancer: his4-S8 MAT**a***	Dershowitz *et al*. 2007
yTAK037	Δ|(-2)|-7B-YUNI300 *MAT**a** CAN1 his7-2 leu2-Δ::kanMX ura3-Δ trp1-289 ade2-1 lys2-ΔGG2899-2900 pol2-M644G agp1::URA3-OR1*	Clausen *et al.* 2015
yTAK066	Δ|(-2)|-7B-YUNI300 *MAT**a** CAN1 his7-2 leu2-Δ::kanMX ura3-Δ trp1-289 ade2-1 lys2-ΔGG2899-2900 pol3-L612G*	Clausen *et al.* 2015
BY4741	*W303 MAT**a** his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	Brachmann *et al*. 1998
VGY129	W303 *MAT**a** ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars609::ars609 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars610::URA3-ars610 BUD4 RAD5*	A. Donaldson lab
2369	Y55 *MATα ura2-1 tyr1-1*	E. Louis lab
2370	Y55 *MAT**a** ura2-1 tyr1-1*	E. Louis lab
45-1	*MAT**a** leu2-3 112 ura3-52 ade2-1 lys2-801 cdc45-1*	C. Nieduszynski lab

CNY167	*MAT**a** ade2-1 his3-11,15 leu2-3 112 trp1-1 ura3-1 can1-100 Gal+ orc5-1*	C. Nieduszynski lab
AUY080	*MAT**a** ade2-1 trp1-1 can1-100 leu2-3 112 his3-11,15 ura3 GAL+ ssd1,d2 RAD5 orc2-1*	C. Nieduszynski lab
CAY109	W303 *MATα ade2-1 his3-11,15 leu2-3 112 trp1-1 ura3-1 can1-100 Gal+ psi+ ars606::ars606 ACS->XhoI ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI*	C. Nieduszynski lab
CAY110	W303 *MAT**a** ade2-1 his3-11,15 leu2-3 112 trp1-1 ura3-1 can1-100 Gal+ psi+ ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI*	C. Nieduszynski lab
DBY72	W303 *MATα RAD5 BUD4 leu2 ura3 trp1 ade2 his3*	C. Nieduszynski lab
AM337	W303 *MAT**a** RAD5 BUD4 leu2 ura3 trp1 ade2 his3 ars609 ars610::URA3-ars610*	C. Nieduszynski lab
EHY69	W303 *MATα pol2-M644G leu2 ura3 trp1 ade2 his3 RAD5 BUD4*	C. Nieduszynski lab
ACY001	W303 *MATα rts1Δ::kanMX*	This Study
ACY004	W303 *MAT**a** rts1Δ::kanMX*	This Study
ACY007	W303 *MATα rif1Δ::kanMX*	This Study
ACY010	W303 *MAT**a** rif1Δ::kanMX*	This Study
ACY036	W303 *MATα/MAT**a** orc2-1/ORC2 rts1Δ::kanMX/RTS1*	This Study
ACY113	W303 *MATα orc2-1 rts1Δ::kanMX*	This Study
ACY044	W303 *MATα/MAT**a** orc2-1/ORC2 rif1Δ::kanMX/RIF1*	This Study
ACY100	W303 *MATα orc2-1 rif1Δ::kanMX*	This Study

ACY079	W303 *MATα/MAT**a** cdc6-1/CDC6 rts1Δ::kanMX/RTS1*	This Study
ACY112	W303 *MAT**a** cdc6-1 rts1Δ::kanMX*	This Study
ACY081	W303 *MATα/MAT**a** cdc6-1/CDC6 rif1Δ::kanMX/RIF1*	This Study
ACY148	W303 *MAT**a** cdc6-1 rif1Δ::kanMX*	This Study
ACY035	W303 *MATα/MAT**a** cdc7-4/CDC7 rts1Δ::kanMX/RTS1*	This Study
ACY096	W303 *MAT**a** cdc7-4 rts1Δ::kanMX*	This Study
ACY042	W303 *MATα/MAT**a** cdc7-4/CDC7 rif1Δ::kanMX/RIF1*	This Study
ACY139	W303 *MAT**a** cdc7-4 rif1Δ::kanMX*	This Study
ACY071	W303 *MATα/MAT**a** dbf4-1/DBF4 rts1Δ::kanMX/RTS1*	This Study
ACY106	W303 *MATα dbf4-1 rts1Δ::kanMX*	This Study
ACY073	W303 *MATα/MAT**a** dbf4-1/DBF4 rif1Δ::kanMX/RIF1*	This Study
ACY104	W303 *MATα dbf4-1 rif1Δ::kanMX*	This Study
ACY031	W303 *MATα/MAT**a** cdc45-27/CDC45 rts1Δ::kanMX/RTS1*	This Study
ACY093	W303 *MATα cdc45-27 rts1Δ::kanMX*	This Study
ACY037	W303 *MATα/MAT**a** cdc45-27/CDC45 rif1Δ::kanMX/RIF1*	This Study
ACY142	W303 *MAT**a** cdc45-27 rif1Δ::kanMX*	This Study
ACY019	W303 *MATα/MAT**a** cdc45-1/CDC45 rts1Δ::kanMX/RTS1*	This Study
ACY087	W303 *MAT**a** cdc45-1 rts1Δ::kanMX*	This Study

ACY051	W303 *MATα/MAT**a** cdc45-1/CDC45 rif1Δ::kanMX/RIF1*	This Study
ACY145	W303 *MAT**a** cdc45-1 rif1Δ::kanMX*	This Study
ACY025	W303 *MATα/MAT**a** cdc9-1/CDC9 rts1Δ::kanMX/RTS1*	This Study
ACY067	W303 *MATα cdc9-1 rts1Δ::kanMX*	This Study
ACY050	W303 *MATα/MAT**a** cdc9-1/CDC9 rif1Δ::kanMX/RIF1*	This Study
ACY120	W303 *MATα cdc9-1 rif1Δ::kanMX*	This Study
ACY069	W303 *MATα/MAT**a** dpb11-24/DPB11 rts1Δ::kanMX/RTS1*	This Study
ACY087	W303 *MAT**a** dpb11-24 rts1Δ::kanMX*	This Study
ACY046	W303 *MATα/MAT**a** dpb11-24/DPB11 rif1Δ::kanMX/RIF1*	This Study
ACY123	W303 *MATα dpb11-24 rif1Δ::kanMX*	This Study
ACY162	W303 *MAT**a** ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 BUD4 RAD5*	This Study
ACY165	W303 *MATα ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 leu2 RAD5 BUD4*	This Study
ACY169	W303 *MATα ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 rnh201::KanMX RAD5 BUD4*	This Study

ACY180	W303/Δ|(-2)|-7B-YUNI300 (1:1) *MAT**a** pol2-M644G; ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610*	This Study
ACY189	W303/Δ|(-2)|-7B-YUNI300 (3:1) *MAT**a** ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 Δrnh201::kanMX pol2-M644G*	This Study
ACY190	W303/Δ|(-2)|-7B-YUNI300 (3:1) *MAT**a** ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 pol2-M644G*	This Study
ACY196	W303/Δ|(-2)|-7B-YUNI300 (7:1) *MAT**a** ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 pol2-M644G rnh201::KanMX*	This Study
ACY200	W303/Δ|(-2)|-7B-YUNI300 (1:1) *MAT**a** pol3-L612G; ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610*	This Study
ACY204	W303/Δ|(-2)|-7B-YUNI300 (3:1) *MAT**a** ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 Δrnh201::kanMX pol3-L612G*	This Study

ACY210	W303/Δ|(-2)|-7B-YUNI300 (3:1) *MAT**a** ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 pol3-L612G*	This Study
ACY217	W303/Δ|(-2)|-7B-YUNI300 (7:1) *MAT**a** ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 pol3-L612G Δrnh201::KanMX*	This Study
ACY221	W303 *MAT**a** ars606::ars606 ACS->XhoI ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 BUD4 RAD5*	This Study
ACY223	W303 *MATα ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 bud4 RAD5*	This Study
ACY226	W303 *MATα ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 BUD4 RAD5*	This Study
ACY237	W303 *MAT**a** pol3-L612G leu2 ura3 trp1 ade2 his3 RAD5 BUD4*	This Study
ACY238	W303 *MATα pol3-L612G leu2 ura3 trp1 ade2 his3 RAD5 BUD4*	This Study

ACY239	W303 *MATα Δrnh201::kanMX ars606::ars606ACS->XhoI ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 BUD4 RAD5*	This Study
ACY240	W303 *MAT**a** Δrnh201::kanMX ars606::ars606ACS->XhoI ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 BUD4 RAD5*	This Study
ACY247	W303 *MAT**a** pol2-M644G ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 BUD4 RAD5*	This Study
ACY249	W303 *MATα pol2-M644G ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 BUD4 RAD5*	This Study
ACY251	W303 *MATα pol3-L612G ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 BUD4 RAD5*	This Study
ACY253	W303 *MAT**a** pol3-L612G ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 BUD4 RAD5*	This Study

ACY258	W303 *MATα pol2-M644G ars606::ars606ACS->XhoI ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 Δrnh201::kanMX BUD4 RAD5*	This Study
ACY259	W303 *MAT**a** pol2-M644G ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 Δrnh201::kanMX BUD4 RAD5*	This Study
ACY261	W303 *MAT**a** pol3-L612G ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 Δrnh201::kanMX BUD4 RAD5*	This Study
ACY264	W303 *MATα pol3-L612G ars606::ars606ACS->XhoI ars607::ars607 ACS->XhoI ars608::ars608 ACS->XhoI ars608.5::ars608.5 ACS->XhoI ars609::ars609 ACS->XhoI ars610::URA3-ars610 Δrnh201::kanMX BUD4 RAD5*	This Study

***Table 2.4. Plasmids used in this study*.**

PLASMID	*DESCRIPTION*	*SOURCE*
YCp-lacZ	Plasmid vector lacking a yeast replication origin. Used for *ARS* assays.	C. Nieduszynski lab
pCA008	YCp-lacZ with *ARS606* inserted.	C. Nieduszynski lab
pCA010	YCp-lacZ with *ARS606::ars606* inserted.	C. Nieduszynski lab
pCA011	YCp-lacZ with *ARS607* inserted.	C. Nieduszynski lab
pCA012	YCp-lacZ with *ARS607::ars607* inserted.	C. Nieduszynski lab
pCA279	Plasmid used for CRISPR to mutate *POL3* to *pol3-L612G*	C. Nieduszynski lab
pAC026	YCp-lacZ with *ARS608.5* inserted.	This study
pAC027	YCp-lacZ with *ARS608.5::ars608.5* inserted.	This study

Table 2.5. Oligonucleotides used in this study

AMPLIFIED REGION	*5' PRIMER: **NAME** SEQUENCE*	*3' PRIMER: **NAME** SEQUENCE*
RIF1Δ::KanMX	**AC0003** TTTTCAGTTCTTTGTGTTTTTCCTC	**AC0004** TGATCCTTTAGAATGGAGAAGATTG
RTS1Δ::KanMX	**AC0005** TAAACCATCGTCGCCGTAA	**AC0006** GGAAGAAGGAAAGCGAAAAGA
TRP1Δ::KanMX	**AC0001** AGAGACCAATCAGTAAAAATCAACG	**AC0002** GCGAAAAGACGATAAATACAAGAAA
5' insertion of RIF1Δ::KanMX	**CA1118** CCATTACGCTCGTCATCAAA	**AC0010** AAGAAACAAGAAGTCAACAGAAGG
3' insertion of RIF1Δ::KanMX	**CA1117** GATAATGTCGGGCAATCAGG	**AC0009** GCGGTAGCATTTCCATCATAA
5' insertion of RTS1Δ::KanMX	**CA1118** CCATTACGCTCGTCATCAAA	**AC0011** GGCATGTCAATACGTCTCGTT
3' insertion of RTS1Δ::KanMX	**CA1117** GATAATGTCGGGCAATCAGG	**AC0012** GGCAAGGTTTACGGAAAAGA
5' insertion of TRP1Δ::KanMX	**CA1118** CCATTACGCTCGTCATCAAA	**AC0007** GAGCTGACAGGGAAATGGTC

3' insertion of TRP1Δ::KanMX	**CA1117** GATAATGTCGGGCAATCAGG	**AC0008** TGTTTGATATTACTGTCAGCGTAGA
ars607 *by ddPCR*	**AC0041** CCATTAGAGACAGAGAAACTATTCATTG	**AC0039** CTTGATCCCAACGTAACTAAATCTC
ARS607 *by ddPCR*	**AC0041** CCATTAGAGACAGAGAAACTATTCATTG	**AC0040** ACTTGATCCCAACGTAACTAAATATAAAC
ars608 *by ddPCR*	**AC0031** GGATGCCGCAATTGAAGAATTA	**CA026** GCGCATAGAAGAAAACTCGAG
ARS608 *by ddPCR*	**AC0031** GGATGCCGCAATTGAAGAATTA	**AC0042** CTTGCGCATAGAAGAAAACTAAAAGT
ars609 *by ddPCR*	**AC0014** GCCAATTCCGGAAAAACTCGAG	**AC0032** GCCGAGAGCTGACAATTTAGTA
ARS609 *by ddPCR*	**AC0013** GAATTTAGCCAATTCCGGAAAAAACAT	**AC0032** GCCGAGAGCTGACAATTTAGTA
ars610 *by ddPCR*	**AC0035** GGATGCACAGTTCAGAGTTATCTA	**AC0033** TCTTCCTTCTGTTCGGAGATTACC
ARS610 *by ddPCR*	**AC0035** GGATGCACAGTTCAGAGTTATCTA	**AC0034** CCACCTAATAATCACCGTTAAACTC

Early control (ChrIV) for ddPCR	**DB013** TCATCTCAACGAGCATGTGC	**DB014** CCTTGTATTAATAGCGACGTTGTC
Late control (ChrIV) for ddPCR	**DB064** ATCCTGCGCGTTGACATAA	**DB065** AGATTCCGTTGCTGGCTATC
Mating type MATα	**CA377** AGTCACATCAAGATCGTTTATGG	**CA379** ACTCCACTTCAAGTAAGAGTTTG
Mating type MAT**a**	**CA377** AGTCACATCAAGATCGTTTATGG	**CA378** GCACGGAATATGGGACTACTTCG
ars606 *specific*	**CN2031** GAAATTGCAAACGAACTCGAG	**CN2036** GCGCAGACGTATCAAACCTAA
ARS606 *specific*	**CN2032** GCGAAATTGCAAACGAAAATATA	**CN2036** GCGCAGACGTATCAAACCTAA
ars607 *specific*	**DB85** TTTAGCTGGGTTTATGGGAGG	**AC0039** CTTGATCCCAACGTAACTAAATCTC
ARS607 *specific*	**DB85** TTTAGCTGGGTTTATGGGAGG	**AC0040** ACTTGATCCCAACGTAACTAAATATAAAC
ars608 *specific*	**CA053** TTGTTGTTGATTGGGTGGAT	**CA026** GCGCATAGAAGAAAACTCGAG

ARS608 *specific*	**CA053** TTGTTGTTGATTGGGTGGA**T**	**AC0042** CTTGCGCATAGAAGAAAACT AAAAGT
ars608.5 *specific*	**AC0079** TTATTAGCTAGGATTATACA CTCGAG	**AC0077** GAATTGATCGGTACCAGAG C
ARS608.5 *specific*	**AC0078** ATGTTTATTAGCTAGGATTAT ACACAT	**AC0077** GAATTGATCGGTACCAGAG C
ars609 *specific*	**AC0014** GCCAATTCCGGAAAAACTCG AG	**AC0015** ACATACTTCTGTGCCCATAG C
ARS609 *specific*	**AC0013** GAATTTAGCCAATTCCGGAA AAAACAT	**AC0015** ACATACTTCTGTGCCCATAG C
ars610 *specific*	**AC0016** TCGGAGAGTTAACAAGCGG C	**AC0033** TCTTCCTTCTGTTCGGAGAT TACC
ARS610 *specific*	**AC0016** TCGGAGAGTTAACAAGCGG C	**AC0034** CCACCTAATAATCACCGTTA AACTC
Spanning ARS606	**CN2035** CCCACTTTGGAAGTGTTTGA A	**CN2036** GCGCAGACGTATCAAACCTA A
Spanning ARS607	**DB85** TTTAGCTGGGTTTATGGGAG G	**CN2014** TGTTAATTCTCAAAAATTCAG GTGA

Spanning ARS608	**CA053** TTGTTGTTGATTGGGTGGAT	**CA023** TTCGACCAGCGCCTATGTAT
Spanning ARS608.5	**AC0075** TGCAATCTGGGTCGCAGAAT	**AC0076** AAGACCACTCTGCCAGAACG
Spanning ARS609	**CN2005** TGCTATACTGAGATCTGGCCAGGGGCA	**AC0015** ACATACTTCTGTGCCCATAGC
Spanning ARS610	**AC0016** TCGGAGAGTTAACAAGCGGC	**AC0017** TCAGCGGCTCCAAATATTGT
POL2 *specific*	**AC0038** CGATGCAGTCGCTACGTATTA	**AC0036** TTTGTAGTCATGATGTTTGGGTACA
pol2-M644G *specific*	**AC0038** CGATGCAGTCGCTACGTATTA	**AC0080** TTTGTAGTCATGATGTTTGGGTAAC
POL3 *specific*	**AC0081** TTGCAACTTTGGATTTCAATTCTTT	**AC0045** CGGAATCAGTGTCACCGTAA
pol3-L612G *specific*	**AC0044** GCAACTTTGGATTTCAATTCTGG	**AC0045** CGGAATCAGTGTCACCGT
RHH201Δ::KanMX	**AC0019** ACAAAGGCAGACATAGTACGCTAAT	**AC0018** AATTGACAGGAAACAAAACTGAGAC

5' insertion of RNH201Δ::KanMX	**AC0020** CCCTTAGCCAAGACGGATCC	**ALU29** CGTCAAGACTGTCAAGGAGG
3' insertion of RNH201Δ::KanMX	**CN2195** TGGTCGCTATACTGCTGTCG	**AC0021** TGGCACATCAGCTTTGTCCT

2.4 Molecular genetic methods

All centrifugation steps for 1.5 ml Eppendorf tubes used a scientific lab supplies Eppendorf 5424R centrifuge, with FA-45-24-11 rotor. Both the non-refrigerated model (~23 °C) and refrigerated model (4 °C) were used. All centrifugation steps for 96-well plates and falcon tubes (15 ml or 50 ml) used a ThermoScientific Heraeus Multifuge X3R centrifuge with TX-750 rotor, set to either 23 °C or 4 °C.

2.4.1 Chemical transformation of E. coli *cells*

Competent DH5α *E. coli* cells were stored at -80 °C, then defrosted on ice immediately prior to transformation. Plasmid DNA (50 – 200 ng) in 5 µl ddH_2O was added to 10 µl cells and incubated on ice for 30 minutes. Cells were heat-shocked at 42 °C for 45 seconds then returned to ice for 2 minutes. 150 µl LB broth was added, then cells were incubated at 37 °C for 30 minutes before the total volume of 165 µl was pipetted onto an LB agar plate with ampicillin and spread using sterile glass beads. Plates were incubated at 37 °C for 24 hours, at which time individual colonies could be selected.

2.4.2 Plasmid extraction from E. coli *cells*

E. coli cells containing a desired plasmid were inoculated into 3 ml LB medium with a selective antibiotic (usually ampicillin), and incubated shaking (200 rpm) at 37 °C for 12 – 16 hours. Cells were centrifuged (3000 rpm, 5 min, RT) and the cell pellet was processed using a Machery-Nagel NucleoSpin plasmid kit, according to the manufacturer's instructions. Briefly, a series of buffers were added to cause cell lysis, clarify the lysate then bind the extracted DNA to a silica

membrane which was washed. Finally, DNA was eluted from the membrane in 50 µl ddH_2O.

2.4.3 DNA quantification

2.4.3.1 Determining DNA concentration

Plasmid and PCR DNA concentration was determined using a DeNovix DS-11 FX spectrophotometer and fluorometer. Concentration of double-stranded genomic yeast DNA was determined using a Qubit (Invitrogen, Q32857) fluorometer following the High-Sensitivity (HS) DNA protocol, according to manufacturer's instructions. Briefly, HS buffer and Qubit dye were mixed in a 199:1 ratio. 1 – 10 µl of DNA (depending on concentration) was added to 199 – 190 µl of this mix, to make a total volume of 200 µl. The Qubit machine was normalised using known standards of 0 ng/µl and 10 ng/µl DNA, then each sample was read in quadruplet and the mean concentration found. Single-stranded DNA was quantified using a NanoDrop™ 2000 (ThermoFisher Scientific) with the type setting 'ssDNA-33'.

2.4.3.2 Determining DNA fragment sizes

DNA fragment length was determined using an Agilent 4200 Tapestation with a High-Sensitivity (HSD1000) ScreenTape which has a size range of 35 – 1000 bp, according to manufacturer's instructions. This service was provided by the University of Oxford Zoology Sequencing team.

2.4.3.3 Quantitative PCR (qPCR)

Quantitative PCR (qPCR) was used to find the concentrations of DNA libraries prior to sequencing. Samples were prepared using the NEBNext Library Quant kit, according to the manufacturer's recommendation. Briefly, DNA libraries were diluted to 1:10000 or 1:100000 using dilution buffer provided by NEB and 4 µl of diluted library was added to 16 µl of a master mix with added primer mix. Triplicates were prepared of each sample, alongside four known standards and a no DNA control. The qPCR programme (95 °C for 60 seconds followed by 35 cycles of 95 °C for 15 seconds and 63 °C for 45 seconds) was run in a Rotor-Gene Q (Qiagen) qPCR machine. The Q-Rex (Qiagen) software was used to generate Ct values and identify anomalous samples. A standard curve was made using the Ct values from the four known standards. This allowed sample DNA concentration to be calculated by adjusting the average concentration from the sample repeats for the average fragment size (known from Tapestation, see 2.4.3.2) and multiplying by the dilution factor of 10000 or 100000 using the formula:

$$\frac{(average\ sample\ concentration\ x\ 399)}{average\ library\ fragment\ size} \times dilution\ factor$$

2.4.4 Polymerase chain reaction (PCR)

PCR reactions were performed in reaction volumes between 10 µl, to verify genotypes of yeast strains (Table 2.3) and plasmids (Table 2.4), and 150 µl to amplify DNA fragments for subsequent use. A typical PCR reaction of 10 µl contained a final volume of 1 µM of each DNA oligonucleotide (one forward and one reverse, Table 2.5), 2 mM magnesium chloride (Bioline, BIO-37026), 1x

BIOTAQ™ buffer (Bioline), 10 – 200 ng template DNA, 0.25 mM mixed dNTPS (dATP, dGTP, dCTP and dTTP) and 1 unit of BIOTAQ™ DNA polymerase (Bioline, 21040; always added last). A standard PCR programme comprised 5 minutes denaturing at 94 °C followed by 30 cycles of 94 °C for 30 seconds, primer annealing at 55 °C for 30 seconds and 1 minute per kilobase DNA extension at 72 °C before a final synthesis step for 5 minutes at 72 °C. PCR programmes were run in a BioRad DNA Engine Thermal Cycler, and completed PCR reactions were stored at 4 °C.

2.4.5 Restriction enzyme digestion of DNA

DNA digests of linear PCR fragments or plasmid DNA were performed in reaction volumes between 10 µl (verification of extracted plasmids) and 50 µl (isolating required DNA fragments for subsequent use). Digests were carried out using enzymes and buffers according to manufacturer's recommendations. Two to 10 units of enzyme were used to digest a maximum of 200 ng of DNA per 10 µl of reaction.

2.4.6 DNA ligation

2.4.6.1 Ligation of PCR products into pGEM®-T Easy

PCR products synthesised using *Taq* polymerase leave a 5' adenine overhang. The pGEM®-T Easy kit (Promega) supplies a vector with a 3' thymidine overhang, allowing high efficiency ligations. Ligations were performed using the pGEM-T Easy kit, according to manufacturer's recommendations modified to use half the reaction volumes.

2.4.6.2 Ligation of restriction enzyme-restricted DNA

To ligate DNA fragments into other vectors, such as YCp-lacZ, both the DNA fragment and plasmid vector were digested using restriction enzymes as described in section 2.4.5, in order to create 3' and 5' overhangs which facilitated ligation. Ligations were carried out in 10 - 50 µl reaction volumes using T4 DNA ligase (NEB) according to manufacturer's recommendations. Reactions were incubated at room temperature for 1 – 2 hours or overnight at 4 °C.

2.4.7 Agarose gel electrophoresis

Gel electrophoresis was used to separate restriction enzyme-digested DNA fragments of different lengths, and to confirm PCR product presence and length. Agarose gels were typically 1.4 % agarose (Lonza SeaKem LE) in 1x TBE buffer (10x buffer contains 108 g/l Trizma base; Sigma, 55 g/l Boric acid; Fisher, 40 ml/l 0.5 M EDTA pH8; VWR) although lower percentage agarose gels (0.9 %) were used for larger DNA fragments in the 1500 – 10000 bp range. All gels had 0.1 % RedSafe (ChemBio) dye added to allow visualisation. 100 ng – 1 µg DNA had 1x orange loading dye added (40 % (w/v) glycerol, 0.5 % SDS, 10 mM EDTA, 0.1 % (w/v) Orange G Sodium Salt; Sigma), before being loaded into the gel wells. Each gel had one lane with 0.5 µg NEB 2-log DNA ladder (N3200S) loaded for comparison. Gels were run in an electrophoresis tank at 70 – 120 V for 20 minutes – 1.5 hours until DNA was sufficiently separated, then visualised on a UVP BioDoc-It Imaging System UV transilluminator.

2.4.8 DNA extraction from agarose gels

Agarose gels with empty lanes between samples were used to isolate DNA fragments. Gels were viewed on a transilluminator and sections containing the desired DNA fragment size were excised using a sterile scalpel and processed using a QIAquick gel extraction kit, according to the manufacturer's instructions. Briefly, the gel was dissolved in buffer at 50 °C, isopropanol was added then the mixture was passed through a QIAquick spin column to bind DNA to a silica membrane. After washing with buffer, the DNA was eluted into 50 µl ddH_2O.

2.4.9 Sanger DNA sequencing

Source BioScience Sanger sequencing service (Oxford branch) was used to confirm PCR product and plasmid sequences. 5 µl of sample (10 ng/µl for PCR products, 100 ng/µl for plasmids) was submitted per reaction with either 5 µl primer (3.2 pmol/µl), or using standard M13 forward and reverse primers from Source BioScience. Plasmid sequences were normally verified twice, in both a forward and reverse orientation.

2.4.10 Ethanol precipitation of DNA

In order to reduce the volume or change the solution of a DNA sample, ethanol precipitation was performed. A volume of 100 % ethanol (typically 2x the volume of the sample) was added with Sodium Chloride (final concentration 0.2 M). The sample was mixed by inversion, incubated at -20 °C for 1 – 2 hours then centrifuged at 4 °C for 15 minutes, 15000 rpm. The supernatant was carefully removed, leaving a DNA pellet. The pellet was washed in 70 % ethanol then resuspended in ddH_2O (typically 50 – 200 µl).

2.4.11 Droplet Digital PCR (ddPCR)

ddPCR was used to quantitatively measure relative copy number of yeast genomic loci. Template DNA was digested into fragments using a restriction enzyme that does no cut within the site of amplification. Digested DNA was processed following manufacturer's recommendations for the QX200 Droplet Digital PCR system (Bio-Rad). Briefly, 0.5 ng of digested DNA in 9 µl H_2O was added to 11 µl EvaGreen Digital PCR SuperMix (Bio-Rad) alongside 2 µl of primers (each 2.5 µM, Table 2.5). To minimise pipetting errors, master mixes were used wherever possible. Prior to PCR, the reaction mix was separated into droplets using oil supplied by Bio-Rad in a droplet generator (Bio-Rad QX200). Droplets were optimised to contain, on average, less than one PCR reaction (DNA and primers). This generated a binary outcome of droplets which were positive or negative for PCR product. The PCR programme was run in a C1000 Touch Thermal Cycler (Bio-Rad) and consisted of 5 minutes denaturing at 94 °C followed by 40 cycles of 94 °C for 30 seconds then primer annealing and synthesis at 60 °C for 30 seconds. After PCR, concentration of DNA in arbitrary units was determined by a droplet reader (Bio-Rad QX200) based on the number of positive vs negative droplets. Mean concentrations of DNA in arbitrary units, from three technical repeats, were generated, alongside maximum and minimum error values, by Bio-Rad Quantasoft software. These values were normalised to early and late controls and given a normalised copy number as follows:

$N_S = \frac{C_S - C_L}{C_E - C_L} + 1$ (S phase cells), $N_S = \frac{C_S - C_A}{C_A} + 1$ (G2 phase cells)

Where N_S was the normalised copy number of sample, and C_S, C_L, C_E and C_A were the concentrations (in arbitrary units) of the sample, late and early controls and the average of the early and late controls, respectively. Diploid regions had their values halved before transformation. Maximum and minimum error values were transformed in the same manner.

2.4.12 Preparation of polymerase-usage sequencing (Pu-seq) DNA

2.4.12.1 Alkali treatment of DNA

Yeast genomic DNA collected from Pu-seq strains contained incorporated ribonucleotides which allowed DNA fragmentation by alkali treatment. DNA was treated according to an established Pu-seq protocol (Keszthelyi *et al.* 2015). Briefly, 15 – 20 µg DNA in 1x TE was incubated at 55 °C for two hours with fresh 5 M NaOH (final concentration 0.3 M).

2.4.12.2 Size selection of DNA

One of two methods for size selection of the resulting ssDNA fragments were used to isolate fragments of the correct size for library making.

2.4.12.2a Size selection of DNA by Blue Pippin

Immediately after alkali treatment, 10 ng of DNA (in 30 µl alkali solution) was loaded onto a Blue Pippin (Sage Science) machine using a 2 % DF agarose cassette (BEF2010) with external marker M1. Blue Pippin software v.5.8 was used to select the range of DNA fragment sizes (300 – 600 bp) and the programme was run according to the manufacturer's recommendations. DNA

was eluded by the Blue Pippin in 40 µl Tris-TAPS (*N*-[Tris(hydroxymethyl)methyl]-3-aminopropanesulfonic acid) buffer and quantified using NanoDrop.

2.4.12.2b Size selection of DNA by gel electrophoresis

Immediately after alkali treatment, 10 ng of DNA (in 50 µl alkali solution) was loaded onto a 0.5x TBE 2 % agarose electrophoresis gel and run alongside two DNA ladders for 2 hrs at 100 V, as outlined in the published protocol (Keszthelyi *et al.* 2015). Briefly, the gel was stained in a 0.5 µg/ml solution of acridine orange solution (ThermoFisher Scientific) then de-stained overnight in ddH_2O. Fragmented ssDNA was observed on a transilluminator as an orange smear. DNA fragments sized 300 – 600 bp (determined by the DNA ladders) were excised using a sterile scalpel and DNA was extracted using a MACHERY-NAGEL NucleoSpin Gel and PCR Clean-up, according to the manufacturer's recommendations. DNA was eluded into 30 µl warm (55 °C) NE buffer (MACHERY NAGEL) and quantified using NanoDrop.

2.4.12.3 Second strand synthesis of ssDNA

In order to make DNA libraries, size-selected dsDNA is required. Therefore a second strand synthesis reaction was performed on the size selected ssDNA from either Blue Pippin or gel electrophoresis, as described in (Keszthelyi *et al.* 2015). Briefly, 100 – 150 ng DNA was incubated with random 8-mer primers, dNTPs (dATP, dCTP, dGTP and dUTP) and T4 polymerase leading to second strand synthesis where dUTP was incorporated into the newly synthesised strand instead of dTTP. 0.5 M EDTA was added to stop the reaction after 20 minutes. Newly synthesised dsDNA was cleaned up using Agencourt AMPure XP beads

(Beckman Coulter) warmed to room temperature, eluded into 58 µl ddH_2O and checked by Tapestation (2.4.3.2) and Qubit (2.4.3.3) to confirm the fragment size range and concentration. DNA was stored at -20 °C.

2.4.13 Making DNA NextSeq libraries

DNA libraries were prepared for sequencing using yeast genomic DNA from synchronous culture experiments, cells sorted by fluorescence activated cell sorting (Sort-seq) and polymerase-usage sequencing (Pu-seq). For the former two applications, 20 ng – 1 µg of extracted DNA in 400 µl 1x TE (10 mM Tris HCl pH 8.0, 1 mM EDTA pH 8.0) was sonicated in a Bioruptor (Diagenode) at 4 °C (20 min, low setting, 30 sec on / 30 sec off). The sonicated DNA was ethanol precipitated using 800 µl ethanol, and resuspended in 25 µl 1x TE.

2.4.13.1 Making DNA libraries from sonicated DNA

DNA fragment sizes were checked by Tapestation (2.4.3.2) to confirm most fragments were in the range 100 – 550 bp. 25 µl DNA in 1x TE from sonication was used as input in NEBNext Ultra II library prep kit (Illumina), according to slightly modified manufacturer's recommendations using half volumes. Briefly, DNA went through an end-prep step followed by adaptor ligation for 25 minutes prior to incubation with USER enzyme. DNA was cleaned up using AMPure XP beads without size selection, then eluted into Tris HCl (10 mM, pH 8.0). A PCR enrichment step was performed, with the number of PCR cycles dependant on the amount of input DNA (6 cycles for 60 ng, 8 cycles for 20 ng, 9 cycles for 16 ng). The PCR enrichment also added universal and indexing primers (NEB Multiplex Oligos kits Set1/Set 2, E7335/E7500), which allowed subsequent

identification of the DNA during next generation sequencing. The final AMPure bead clean up step was performed, again using the manufacturer's recommendations with half volumes. However, DNA was eluted after clean-up into 26 µl 0.1x TE, 25 µl of which was used to fully repeat the clean-up process before eluting into 33 µl 0.1x TE. Supernatant was carefully removed from the AMPure XP beads resulting in 30 µl final library, as recommended by the manufacturer. The finished libraries were checked by Tapestation to ensure fragment sizes were still predominantly in the 100 – 550 bp range (this size increased slightly relative to post-sonication due to adaptor and index addition), and by qPCR to find a precise concentration. DNA libraries were stored at 4 °C for up to one week, then at -20 °C for long term storage. The number of freeze-thaw events was minimised to prevent DNA shearing.

2.4.13.2 Making DNA libraries from Pu-seq DNA

DNA from Pu-seq was also processed using the NEBNext Ultra II library prep kit (Illumina). Samples were processed as in section 2.4.13.1, with some key differences, according to an established Pu-seq protocol (Keszthelyi *et al.* 2015). Fragmented dsDNA from Pu-seq (2.4.12.3) was eluted into ddH_2O and 50 µl of this DNA solution was used as input for the end-prep. The manufacturer's recommended volumes were used as stated, and not halved, during Pu-seq library making. During adaptor ligation, enhancer enzyme was replaced by 1 µl of ddH_2O. Importantly, due to the incorporation of dUTPs in the second strand synthesis of fragmented Pu-seq DNA, the incubation with USER enzyme after adaptor ligation was not performed. DNA from the adaptor ligation reaction (93.5 µl) had ddH_2O added to give a volume to 100 µl. Clean up using AMPure XP

beads was performed with size selection for 250 – 600 bp DNA fragments, using 30 µl beads in the first bead addition, and 40 µl bead in the second bead addition steps. DNA was eluted into 25 µl 0.1x TE pH8.0 and 20 µl was used in the PCR enrichment reaction. The PCR enrichment reaction was modified to include the USER enzyme. This digested the uridine-containing second strand synthesised during preparation of Pu-seq DNA and ensured PCR enrichment of only the strand of interest. This step is crucial to Pu-seq being able to detect strand-specific effects. The modified PCR enrichment reaction consisted of 25 µl NEBNext® Ultra™ II Q5 Master Mix, 2.5 µl indexing primer, 2.5 µl universal primer, 20 µl DNA and 3 µl USER enzyme (from the NEBNext Ultra II kit adaptor ligation step). The PCR reaction was performed according to manufacturer's recommendations and the final clean-up was performed using 50 µl AMPure XP beads to account for the slightly increased volume of the PCR reaction. DNA was eluded after clean up into 26 µl ddH_2O, 25 µl of which was used to fully repeat the clean-up process (using 25 µl AMPure XP beads) before eluting into 23 ddH_2O. Supernatant was carefully removed from the AMPure XP beads resulting in 20 µl final library. Final libraries had their fragment sizes and final concentration confirmed, as in 2.4.12.1 and were stored in the same manner.

2.4.14 NextSeq DNA Sequencing

Due to sequencing technology constraints, multiple libraries were sequenced simultaneously. A pooled library of the total number of samples (n) was created in volume n µl with each library at a concentration of 4 nM. DNA sequencing was performed on a NextSeq 500 next-generation sequencer (Illumina). Samples were prepared immediately prior to loading onto the reagent cartridge according

to the manufacturer's recommendations. Briefly, the 4 nM pooled library was denatured in 0.2 N NaOH (18 µl PW1 buffer provided by Illumina, 2 µl NaOH) and then neutralised using Tris HCl and HT1 buffer (Illumina). The library was further diluted in HT1, producing a final pooled library concentration of 2.2 pM, and was kept on ice. A BaseSpace (Illumina) account was used to set parameters for the sequencing run: these were usually single-end reads with six indexing cycles and 80 read cycles. Paired-end reads were used for polymerase-usage sequencing with six indexing cycles, 80 read 1 cycles and 80 read 2 cycles. Data were downloaded from BaseSpace and uploaded to a university server, then analysed using Unix-based tools (described in 2.6).

2.5 Yeast methods

2.5.1 Growth conditions

Yeast strains were cultured both in liquid and on solid YPAD medium, and manipulated according to established practices (Treco & Winston 2008). For long term storage cells from a single colony were mixed into 25 % glycerol (sterile, 1 ml) and kept at -80 °C.

2.5.1.1 Yeast growth on agar

Strains were recovered in patches or streaked to single colonies on YPAD or selective medium plates. Plates were incubated at the appropriate temperature, usually 30 °C, then stored at 4 °C for up to 4 weeks. Colony transfer to other plates was achieved via re-streaking or replica plating. During replica plating, an agar plate was pressed firmly onto a sterile velvet surface, using a replica stone.

A second plate was pressed firmly onto the same position and incubated at the appropriate temperature, usually 30 °C.

2.5.1.2 Yeast growth in liquid culture

Starter cultures were set up using cells from a single colony inoculated into 5 ml of YPAD or selective medium, grown shaking at 200 rpm for 12 - 16 hrs. Non-temperature-sensitive strains were grown at 30 °C, temperature-sensitive strains were grown at 23 °C. Larger liquid cultures of 15 - 800 ml were used in most yeast methods. These were inoculated using starter cultures and were grown using shaking water baths or incubators with the same parameters as above. Yeast cell density was inferred from OD_{600} readings (using a Thermofisher BioMate3 spectrophotometer) and mid-log phase cells were taken to be those where at least two doublings of the initial OD_{600} reading had occurred and the current OD_{600} reading was 0.5 - 1.0.

2.5.2 Calculating mean generation time

Liquid cultures of cells were inoculated and grown so two doublings of the initial OD_{600} had occurred. OD_{600} cell density readings were then taken at regular intervals, for a minimum of two further OD_{600} doublings within the exponential growth phase. At least six readings were taken to calculate the mean doubling time.

2.5.3 Dilution series assay

Starter yeast cultures were grown in YPAD or selective medium, and their OD_{600} measured. Cells were diluted using the same medium to 10^7 cells/ml, before

serial 10-fold dilutions were prepared. Dilution spots of 5 µl were pipetted onto YPAD or selective medium plates and incubated at stated temperatures. Unless otherwise stated, plates were photographed after two dates and images were analysed qualitatively by observation of relative colony growth.

2.5.4 High efficiency yeast transformation

High efficiency transformation was carried out according to a slightly modified established protocol (Gietz and Shiestl, 2007). Yeast cells were transformed with either plasmid DNA or linear DNA designed with homologous regions for targeted integration to the yeast genome. A starter yeast culture was grown and used to inoculate a larger liquid culture at an OD_{600} of 0.1 – 0.2 (5 ml / reaction). Cells were collected in mid-log phase, centrifuged (3000 rpm, 5 min) then resuspended in 20 ml ddH_2O. Cells were washed with 20 ml ddH_2O then resuspended in 1 ml ddH_2O. For each reaction, 100 µl cells was transferred to a 1.5 ml Eppendorf tube and centrifuged (13000 rpm, 30 s) then 360 µl transformation mix (240 µl PEG 3350 50 % w/v, 36 µl lithium acetate, 50 µl carrier DNA, 34 µl DNA in ddH_2O) was added to the cell pellet and mixed thoroughly by pipetting and vortexing. Cells were incubated at 42 °C for 40 min (or overnight at room temperature for temperature-sensitive strains), then centrifuged (13000 rpm, 30 s) and resuspended in 1 ml ddH_2O. Cells were plated at multiple densities (most commonly 5 % and 95 %) onto YPAD or selective medium plates. G418-resistant cells were plated onto YPAD plates then replica plated onto G418 plates after 24 hrs. Plates were incubated for 3 - 4 days at 30 °C for non-temperature-sensitive strains and 23 °C for temperature-sensitive strains.

2.5.5 ARS assay in yeast

ARS assays were used to determine the origin (ARS) efficiency of genomic or inactivated regions of DNA. A 600 – 1000 bp region of DNA surrounding the genomic or inactivated origin was amplified by PCR (2.4.4). This PCR fragment and the YCp-lacZ plasmid were digested by *Spe*I and *Not*I restriction enzymes (2.4.5) then ligated (2.4.6.1) to form a YCp-lacZ plasmid containing the origin site of interest. 200 ng of plasmid was transformed (2.5.4) into the yeast strain BY4741 (Baker Brachmann *et al.* 1998), which is *ura3Δ0*. This prevented recombination of the *URA3* selective marker on the YCp-lacZ plasmid into the *S. cerevisiae* genome, meaning only cells where the plasmid was retained were able to survive. Cells were plated in three concentrations (1 %, 5 % and 94 %) onto selective –URA plates and colony growth was imaged using a UVP BioDoc-It Imaging System after a certain number of days (usually 2 or 4).

2.5.6 Colony PCR of yeast cells

Single colonies of strains to be tested were selected, and resuspended in 2 µl SPZ buffer (2.5 mg/ml Zymolyase 100T; 30 ml 2M Sorbitol, 4.05 ml 1 M Na_2HPO_4, 0.95 ml $Na_2H_2PO_4$, 15 ml H_2O, pH 7.5), then incubated at 37°C for 30 minutes in a thermocycler. After dilution with 30 µl ddH_2O samples could be frozen at -20 °C or 1 µl used immediately in a 10 µl PCR reaction (Primers shown in Table 2.5).

2.5.7 Mating type assays for yeast

2.5.7.1 Plate-based mating type assay for yeast

Single colonies were replica plated or re-streaked with *MAT**a*** (2679) and *MATα* (2670) tester strains (Table 2.3) then incubated at permissive temperatures. After

24 hours, strains were replica plated onto minimal medium then incubated at the same temperature for a further 48 hour. Growth when mated to either tester strain indicated the original strain was a haploid of the opposite mating type. No growth when mated to either tester strain indicated a diploid, as diploids were unable to mate to the tester strains, which contain essential genes for survival on minimal medium.

2.5.7.2 PCR-based mating type assay for yeast

Alternatively, a colony PCR using mating type tester primers (Table 2.5) for *MAT**a*** and *MATα* alleles was performed to determine yeast mating types.

2.5.8 Mating yeast strains

2.5.8.1 Mating yeast strains on plates

A single colony of one *MAT**a*** and one *MATα* strain were streaked on top of each other onto YPAD or selective medium and mixed, then incubated at a permissive temperature (usually 30 °C) for 24 hours. After 24 hours, strains were streaked to single colonies on YPAD or selective medium then incubated at the same temperature for a further 24 hours. Single diploid colonies were selected using a mating type assay.

2.5.8.2 Mating temperature-sensitive yeast strains in liquid

Some temperature-sensitive strains failed to mate on plates, therefore a liquid mating was used. Starter cultures of both strains were grown and used to co-inoculate 200µl of liquid YPD. The temperature-sensitive strain and the non-temperature-sensitive strain were inoculated in a 10:1 ratio, then incubated at a

permissive temperature, shaking at 200 rpm, for 8-16 hours. Twenty microliters of the mating were plated using glass beads onto YPAD or selective medium, then incubated at a permissive temperature for a further 24 hours. Single diploid colonies were selected using a mating type assay.

2.5.8.3 Sporulation and tetrad dissection

Diploids were streaked onto sporulation medium plates and incubated at room temperature for 5 - 7 days. After this time, tetrads can be confirmed by viewing down a light microscope at 40x objective magnification (400x total magnification). Small numbers of sporulated cells were added to a 20 µl solution comprising 1 µl Zymolyase 20T and 19 µl dissection buffer (0.8 M sorbitol, 10mM EDTA, 10mM $NaPO_4$: pH 7.2). The cells were incubated in a 37°C water bath for 30 minutes, then diluted with 100 µl ddH_2O. A smaller volume (20 - 40 µl) of this cell mixture was pipetted in a thin strip in the centre of YPAD dissection plates. A Singer Instruments Spore Play dissection microscope with a glass needle was used to manipulate tetrads, separating sister colonies. Dissected tetrads were incubated for 2 - 4 days at 30 °C for non-temperature-sensitive strains and 23 °C for temperature-sensitive strains.

2.5.8.4 Backcrossing

In the case of some desired mutations present in non-W303 background yeast, W303 strains were crossed to other *S. cerevisiae* backgrounds in order to obtain the desired mutation. To reduce background genetic mutations, which could potentially alter mean generation time or have other undesired effects, these strains were backcrossed twice to a W303 parental strain.

2.5.9 Harvesting and fixing yeast cells from liquid cultures

2.5.9.1 Large liquid yeast cultures

Yeast cells growing in large liquid culture (> 2 ml) were centrifuged at 3000 rpm for 5 minutes. Medium was removed and cells were washed twice in 5 ml ddH_2O. Cell pellets could either be fixed (resuspended in 20 ml 70 % ethanol) or immediately frozen at -20 °C. Fixed cells were stored at 4 °C for up to a year. Prior to their use, fixed cells were centrifuged using the same conditions from their harvest to make a cell pellet, then washed in ddH_2O.

2.5.9.2 Small liquid yeast cultures

Small liquid cultures of yeast cells (< 2 ml) were centrifuged at 13000 rpm for 5 minutes. Medium was removed and cells were washed twice in 1 ml ddH_2O. Cell pellets could either be fixed (resuspended in 1 ml 70 % ethanol) or immediately frozen at -20 °C. Fixed cells were stored at 4 °C for up to a year. Prior to their use, fixed cells were centrifuged using the same conditions from their harvest to make a cell pellet, then washed in ddH_2O.

2.5.10 Collection of stationary phase yeast cells

A starter culture of yeast cells was grown in YPAD or selective medium at 30 °C. The OD_{600} reading was measured at 48 hrs, 72 hrs and 96 hrs. Cells were harvested and frozen at 96 hrs, or when the OD_{600} reading no longer increased within a 24-hour period, whichever was later.

2.5.11 Arresting yeast cells in alpha-factor

A starter culture of yeast which were *MAT**a*** and wild-type for *BAR1*, was set up in YPAD at 23 °C. This was used to inoculate a larger liquid culture (20 – 100 ml), which was grown overnight to an OD_{600} of 0.18 – 0.20. Alpha factor was added to make a final concentration of 0.45 µM (15 µl of 3 mM alpha factor per 100 ml culture) and the culture was incubated shaking (200 rpm) at 30 °C. Alpha factor was added in the same manner at regular intervals (1/3 the strain's mean generation time) for at least 1.5 mean generation times. Cells were considered arrested when less than 5 % of cells had buds, and shmooing was observed. At arrest cells were either harvested or released.

2.5.12 Synchronous release of yeast cells for synchronous culture experiments

Arrested cells were synchronously released into YPAD by addition of 1 µl pronase (20 mg/ml) per 4 ml culture. Addition of pronase was taken as 'time 0' in synchronous culture experiments. Cell cultures were incubated in a shaking water bath from time 0, to prevent changes in temperature when collecting samples.

2.5.12.1 Collection of synchronous culture samples for flow cytometry

Samples for flow cytometry were collected prior to alpha factor addition (asynchronous), prior to release from alpha factor (time 0) and at regular intervals for the following 2 – 3 hours to capture one synchronous cell cycle after alpha factor release. Flow cytometry samples were collected by sampling 0.5 – 1 ml culture and adding 10 % sodium azide to a final concentration of 0.1 %. Samples were mixed by inversion, then kept on ice until all samples were collected. After

collection of all flow cytometry cell samples, the cells were harvested and fixed in ethanol.

2.5.12.2 Collection of synchronous culture samples for DNA extraction

Samples for DNA were collected prior to release from alpha factor (time 0) and at specified time points during the synchronous cell cycle after alpha factor release. DNA samples were collected by sampling 5 - 10 ml culture and adding this onto a frozen pellet of sodium azide solution (0.2 M EDTA, 0.1 % sodium azide), 1 ml pellet per 10 ml culture. Samples were mixed by inversion to resuspend the pellet, then kept on ice until all samples were collected. After collection of DNA samples, the cells were harvested and the pellets frozen.

2.5.13 'Smash and Grab' DNA extraction of yeast cells

This phenol-chloroform DNA extraction method, based on Rose *et al.* (1990), was used to extract genomic DNA from yeast cell pellets. To the thawed cell pellet, 0.2 g zirconia/silica beads (0.5 mm, Thistle Scientific), 0.2 ml lysis buffer (10 mM Tris pH 8.0, 1mM EDTA, 100mM NaCl, 1 % w/v SDS, 2 % w/v Triton X-100) and 0.2 ml phenol-chloroform-isoamyl alcohol (25:24:1, pH 8.0) were added. The tube was vortexed three times in a FastPrep-24 for one minute each to lyse cells, before centrifugation (12000 rpm, 3 min). The aqueous phase was transferred to a new tube and 0.2 ml 1x TE was added. Tubes were mixed by inversion, centrifuged (12000 rpm, 3 min), and the aqueous phase was transferred to a new tube. After addition of 0.4 ml chloroform, the inversion, centrifugation and transfer steps were repeated. RNAseA (10 µl of a 10 mg/ml solution) was added and samples were incubated at 37 °C for 30 min then proteinase K (10 µl of a 20

mg/ml solution in TE) was added before a further incubation of one hour. The addition, inversion, centrifugation and transfer steps above were repeated for addition of 0.4 ml phenol-chloroform-isoamyl (a phenol chloroform extraction) and two repeats adding 0.4 ml chloroform (two chloroform extractions). Finally, DNA was ethanol precipitated by addition of 0.4 ml ethanol, and resuspended in 50 µl ddH_2O.

2.5.14 Measuring DNA content by flow cytometry

Fixed cells were washed twice in 1 ml sodium citrate (50 mM, 0.45 µm filter sterilised, pH 7.0) then pelleted (13000 rpm, 1 min, RT) and resuspended in 1 ml 50 mM sodium citrate with RNaseA (0.25 mg/ml). Samples were incubated at 50 °C for 1 hour then 20 µl of 20 mg/ml proteinase K was added and samples were incubated for a further hour at 50 °C. Cells were then pelleted (13000 rpm, 1 min, RT) before resuspension in 1 ml 50 mM sodium citrate with 1x SYTOX green (Thermofisher). Cells were stored in the dark either at 4 °C for at least 12 hours, or at room temperature for at least 2 hours before use. Samples were stored with minimal light exposure at 4 °C for up to one week before analysis. Immediately prior to flow cytometry, cells were sonicated with a tip sonicator for 6 - 10 seconds, three times. Flow cytometry was performed on a BD LSR Fortessa X-20 cell analyser (BD Biosciences) which is able to detect SYTOX green. Parameters for forward scatter (FS), side scatter (SS) and fluorescence detector (BL1) were kept consistent between samples within one experiment, but were adjusted slightly for each experiment to optimally detect different strains within the following parameters: FS 580 – 650 V, SS 250 – 305 V, BL1 400 – 430 V.

2.5.15 Fluorescence activated cell sorting (FACS) of yeast

A 25 ml culture of mid-log phase cells was harvested and fixed. Fixed cells were washed twice in 7.5 ml sodium citrate (50 mM, 0.45 µm filter sterilised, pH 7.0) then pelleted (5000 rpm, 5 min, RT) and resuspended in 0.5 ml 50 mM sodium citrate with RNaseA (100 µg/ml). Samples were incubated at 50 °C for 1 hour then 25 µl of 20 mg/ml proteinase K was added and samples were incubated for a further hour at 50 °C. Cells were then pelleted (13000 rpm, 1 minute, RT) before resuspension by vortexing in 5 ml 50 mM sodium citrate with 10x SYTOX green (Thermofisher) and 0.1 % sodium azide. Cells were stored in the dark at 4 °C for at least 12 hours before use. Samples were stored with minimal light exposure at 4 °C for up to four weeks before analysis. Immediately prior to cell sorting, cells volumes no greater than 2 ml were sonicated with a tip sonicator for 6 - 10 seconds, three times. Cells were sorted by the Sir William Dunn School flow cytometry facility using a MoFlo (Coulter Beckman) fluorescence activated cell sorter (FACS), capable of detecting SYTOX green. Samples were sorted into yeast cell cycle stages, collecting 5 – 15 million cells for each sample. The output from the FACS machine was yeast cells in saline sheath fluid at ~1 million cells per millilitre. Ethanol (100 %) was added to these samples (1/3 – 1/4 volume) and samples were vortexed, then stored at 4 °C for no more than 3 days. Cells from FACS were pelleted by sequentially centrifuging (15000 rpm, 20 min, 4 °C) the sample with added 100 % ethanol in an eppendorf tube, then removing the supernatant. The final cell pellet was resuspended in 1 ml ddH_2O. Five microliters of these cells were added to 1 ml 50 mM sodium citrate with 1x SYTOX green and the sample checked by flow cytometry to confirm the correct yeast cell cycle stage had been enriched by FACS. Both the resulting flow cytometry sample and

the cells in ddH_2O were ethanol precipitated by adding 1/3 – 1/4 volume of 100 % ethanol and centrifuging (15000 rpm, 20 min, 4 °C). The resulting cell pellets were combined by resuspension in 100 µl ddH_2O and repeating the ethanol precipitation step.

2.5.16 DNA extraction of cells sorted using FACS

The combined cell pellet from FACS was resuspended in 0.5 ml of 50 mM Tris-HCl and 0.1 mM EDTA. 50 µl Zymolyase (10 mg/ml) and 0.5 µl ß-mercaptoethanol were added and the sample was incubated at 37 °C for 30 minutes. Following incubation, 32 µl of a master mix of 5 µl RNaseA (10 µg/ml), 17 µl NaCl (3 M) and 10 µl proteinase K (20 mg/ml) was added, followed by 25 µl of 10 % SDS solution. The sample was incubated at 55 °C for 1 – 2 hours, then two phenol-chloroform extractions and one chloroform extraction performed, as described in section 2.5.12. Finally, the DNA was ethanol precipitated in 0.5 ml ethanol and resuspended in 200 µl ddH_2O.

2.5.17 Polymerase Usage sequencing (Pu-seq)

Pu-seq was used to examine strand-specific DNA synthesis by Pol2 and Pol3 in *S. cerevisiae*. Experiments were carried out according to an established Pu-seq protocol (Keszthelyi *et al.* 2015), with some modifications to optimise for *S. cerevisiae* cells.

2.5.17.1 Collection of *S. cerevisiae* cells for Pu-seq

A starter culture of yeast which were ∆*rnh201* and *pol2-M644G* or *pol3-L612G*, was set up in YPAD at 30 °C. This was used to inoculate a larger liquid YPAD

culture (10 ml) at OD_{600} 0.05 allowing a mean generation time experiment to be performed after 2 doublings had occurred. The 10 ml culture was grown a further 12 – 16 hours then used to inoculate a 25 ml YPAD culture at OD_{600} 0.05 and a second mean generation time experiment was performed using this culture to confirm the mean generation time. This mean generation time was used to calculate the required starting OD_{600} for a 400 ml YPAD with Ampicillin (Amp) culture which would be OD_{600} 0.8 after 16 – 20 hours. YPAD for 400 ml culture was pre-heated to 30 °C. Once the OD_{600} 0.8 was reached, the culture was split equally into two, with 200 ml YPAD with Amp added to make two 400 ml cultures. After approximately one doubling time (final OD_{600} 0.8 – 1.3) cells from both 400 ml cultures were pooled and pelleted using a JLA 8.1000 rotor in a Beckman Coulter Avanti J-26S XP centrifuge (5 min, 5000 rpm, 4 °C). Supernatant was removed and the cell pellet was washed by resuspension in 35 ml ddH_2O. Cells were moved to a 50 ml Falcon tube and pelleted (5 min, 3000 rpm, 4 °C) before supernatant was removed completely. The pellet was snap-frozen in a bath of dry ice covered with 100 % ethanol, then stored at -80 °C. Pellets were thawed at room temperature before DNA extraction.

2.5.17.2 DNA extraction from cells for Pu-seq

A thawed Pu-seq cell pellet was resuspended in 2 ml NIB buffer (17 % glycerol, 50 mM MOPS (3-(*N*-morpholino)propanesulfonic acid), 150 mM potassium acetate, 40 mM EDTA, adjusted to pH 7.2 using KOH) with 5 mg/ml Zymolyase and incubated at 37 °C for 15 min. The treated cell pellet had DNA extracted as described elsewhere (Keszthelyi *et al.* 2015). Briefly, cell lysis was confirmed using an SDS solution then cells were pelleted and resuspended in 2 ml QIAGEN

buffer G2. Cells were treated sequentially first with 100 µl RNase A (10 mg/ml) and incubated (30 min, 37 °C) then 100 µl fresh proteinase K (Sigma P2308, 20mg/ml) and 100 µl 30 % (wt/vol) *N*-lauroyl sarcosine and incubated (60 min, 55 °C). After centrifugation, the supernatant was isolated and the remaining pellet was re-treated with QIAGEN buffer, RNase A, proteinase K and *N*-lauroyl sarcosine using half the previous volumes. The resulting supernatant was combined with the first supernatant and this volume was applied to an equilibrated QIAGEN 100/G genomic tip, following the manufacturer's recommendations. The resulting DNA pellet was resuspended in 200 µl TE and the DNA quantified. Typical yields from *S. cerevisiae* cells were 450 – 1200 ng/µl which is higher than that observed for *Schizosaccharomyces pombe* (Keszthelyi *et al.* 2015).

2.6 Bioinformatics

2.6.1 Specialised computer programmes

Table 2.6. Scientific computer programmes used in this study.

PROGRAMME	*APPLICATION*
Microsoft Excel	Plotting graphs, statistical analysis, ddPCR and qPCR further data analysis
FlowJo™ v10	Flow cytometry analysis
IGV V2.3.94	Next generation sequencing analysis
QuantaSoft™ (Bio-Rad)	ddPCR programme control and generation of concentration values
Q-Rex (Qiagen)	qPCR programme control and generation of Ct values
FastQC	Quality control for next generation sequencing files

2.6.2 Web tools

The *Saccharomyces cerevisiae* reference genome (sacCer3) was accessed from the UCSC genome browser (genome.ucsc.edu/). Sections of this reference containing regions of interest were downloaded as FASTA files from UCSC and uploaded as DNA files into Benchling (www.benchling.com), an online workspace for sequence alignment, design and visualisation. Benchling was used to design primers, predict PCR product sizes and outcomes of restriction digests for molecular cloning. Primer3 Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) was used to confirm primers fell within quality control parameters and would not form secondary structures. Sanger sequencing files were also uploaded into Benchling for comparison to either the reference genome or modified DNA sequences.

Supplementary information about the *S. cerevisiae* genome, such as genes or *ARS* sequences contained within a region of interest, was obtained from the *Saccharomyces* Genome Database (www.yeastgenome.org) or the DNA Replication Origin Database (OriDB, cerevisiae.oridb.org). Information on gene deletion cassette *S. cerevisiae* strains from the *S. cerevisiae* Genome Deletion Project (SGDP), and the sequence of primers to confirm cassette integration to the correct genomic loci, were found on the SGDP website (www-sequence.stanford.edu/group/yeast_deletion_project/).

Next-generation sequencing run information was input on the online system BaseSpace (basespace.illumina.com/dashboard). The user account was linked to each sequencing run and data were automatically uploaded to the BaseSpace account, available for download. Data were downloaded and formatted into bed files using Unix tools (described in 2.6.3). Bed files for Sort-

Seq and synchronous culture experiments were analysed using the Shiny application plotGenome (dzmitry.shinyapps.io/plotgenome_v2/). plotGenome is an online version of a programming script written in the language 'R', developed by a current lab member, Dzmitry Batrakou.

Table 2.7. Web tools used in this study.

PROGRAMME	*APPLICATION*
Benchling	Plasmid map generation, Sanger sequencing analysis, *In silico* PCR amplification and restriction enzyme digests, primer design
UCSC	Source of *S. cerevisiae* reference genome
Primer3 Plus	Primer design and secondary structure analysis
SGD	Source of *S. cerevisiae* information, particularly genetic interactions
OriDB	Source of information about *S. cerevisiae* DNA replication origins
SGDP website	Source of information about the *S. cerevisiae* Genome Deletion Project yeast strains and primers
BaseSpace (Illumina)	Next generation Illumina sequencing projects, input of settings and data output
plotGenome v2	A shiny app developed by a lab member (Dzmitry Batrakou) used for sort-Seq and synchronous culture experiment data analysis

2.6.3 Unix-based tools

2.6.3.1 Data retrieval and quality control assessment

Unix commands were issued through the terminal to analyse next generation sequencing data. Data were retrieved from BaseSpace using BaseMount (Illumina, v0.15.39.2013) and placed onto a Nieduszynski lab server, where raw

sequencing data is stored long-term and backed up by the University server. Sequencing files were analysed by FastQC using the command: fastqc –t 8 –o location/of/fastqc/app filename.fastq.qz –noextract. This generated html files which could be viewd in a web browser to inspect base statistics, per base quality scores, sequence content, GC content and N content, per sequence quality scores, duplication levels, GC content and length distribution, over represented sequences and Kmer content. Where data were rejected this was typically due to failing the sequence content and GC content per base parameters, indicating that there was an overly high level of 'G' bases. In Illumina sequencing, a 'G' is indicated by a lack of fluorescent dye, so an abundance of 'G' bases over 'C' bases suggests a poor-quality sequencing run.

2.6.3.2 Mapping sequencing reads to the reference genome

Initial sequencing data analysis was done using a shell script created by Prof. Conrad Nieduszynski and updated by Dr. Dzmitry Batrakou, named autoMap.sh. First, the four lane files from paired-end read files were concatenated to create one read one (R1) file and one read two file (R2) using the commands: cat sample_name_L00[1-4]_R1_001.fastq.gz > sample_name_R1.fastq.gz and cat sample_name_L00[1-4]_R2_001.fastq.gz > sample_name_R2.fastq.gz. This step was unnecessary for single end read files, which autoMap could process in the four-lane format. A summary text file (samples_summary.txt) of all files to be analysed was created in the format:

sample_name_1* Sample:brief_name_1

sample_name_2* Sample:brief_name_2…

autoMap was run for single end reads using the command: location/of/autoMap.sh –f samples_summary.txt –g referenceGenome –r single, and for paired end reads the same command was used however the word ‘single’ was substituted with ‘paired’. Briefly, autoMap detected each sample, and for each sample it created a bam and a bed file with uniquely mapped reads: for example, brief_name_1.bam, brief_name_1.bed, brief_name_2.bam and brief_name_2.bed.

2.6.3.3 Confirming genetic modifications to ARS sites

Genetic modification to ARS sites in strains sequenced by next generation sequencing was confirmed using the following samtools command: samtools tview brief_name.bam reference_genome -p location_of_interest.

2.6.3.4 Synchronous culture experiment and Sort-seq sequencing analysis

The bed file sample_name.bed created from autoMap.sh was uploaded into the plotGenome Shiny application. Replicating samples (resulting from S phase enrichment by FACS or after release from alpha factor by pronase) were normalised to non-replicating samples (resulting from G1 or G2 phase enrichment by FACS, stationary cells or alpha factor arrested G1 phase cells). The ratio between replicating and non-replicating samples was automatically generated, then adjusted if necessary to ensure samples had a normalised DNA copy number where un-replicated DNA has a copy number of 1 and replicated DNA has a copy number of 2. Plots of normalised copy number and raw data were downloaded from the plotGenome application.

2.6.3.5 Pu-seq sequencing analysis

The bam files created from autoMap.sh were used in Pu-seq analysis. Complementary yeast strains, one with mutant polymerase epsilon (Polε) and one with mutant polymerase delta (Polδ), were required for Pu-seq analysis. Therefore, a shell script which analysed both strains in parallel was written in collaboration with Phoebe Oldach, a post-graduate student in the Nieduszynski lab: post_mapping_single_strand_file_generation_two_polymerases.sh (Appendix 8.1). This script was a modified version of that used in (Keszthelyi *et al.* 2015).

Briefly, this script accepted a Polε (brief_name_ε.bam) and a Polδ (brief_name_ δ.bam) bam file and used samtools Flag-values to select reads which mapped to the reference genome, mapped in a proper pair (reads one and two are present) and were specific to either the forward (Watson) or reverse (Crick) strand. The Flag-tag also ensured that the read was the first in pair, meaning its 5' base derived from adjacent to the ribonucleotide incorporation site where DNA is fragmented during the preparation of Pu-seq library DNA. A text file was outputted from samtools which summarised the number of ribonucleotide incorporations at each genomic locus by counting the number of 5' reads which ended at the adjacent position (output_polymerase_strand_offset.bg). This process was completed for each polymerase, strand specifically. Secondly bedtools was used to correct the genomic position offset samtools created (samtools starts chromosome numbering at base pair '0', whereas bedtools and downstream application start as base pair '1') This created files in the format (output_polymerase_strand_correct.bg). Bedtools was also used to create a summary file showing every location where one or more ribonucleotides was

incorporated, which strand this was one and which polymerase incorporated it. This summary also included the base (A, C, G or T) present at that site which allowed later analysis of the prevalence with which ribonucleotides were incorporated at each nucleotide (output_complete.tab).

In order to visualise the data, ribonucleotide incorporation was grouped into 300 bp bins. The grouping into 300 bp bins required the use of another file, which gave the bin sizes. This was generated from any bam file mapped to the reference genome using the following commands; the first, samtools, command created a text file (sacCer3_300bp.txt) which listed each chromosome and its length. The second, bedtools, command then used windows of 300 bp to make bins for each chromosome.

```
samtools idxstats brief_name.bam | awk 'BEGIN {OFS="\t"} {if ($2>0) print ($1,$2)}' > sacCer3_300bp.txt
bedtools makewindows -g sacCer3_300bp.txt -w 300 > sacCer3_300bp.bed
```

From the ribonucleotide files in the format output_polymerase_strand_correct.bg and the sacCer3_300bp.bed bins file, three types of file were created sequentially; csv files suitable for use in an existing R application (Keszthelyi *et al.* 2015), wig files suitable for use in IGV, and bed files suitable for import into R software for plotting. Wig files were imported into IGV for visualization.

2.6.3.6 Examining genome coverage

In order to see the level of genome-wide coverage of sequencing reads after mapping the following bedtools command was used: bedtools genomecov –ibam brief_name.bam –bga > brief_name_coverage.bed. This bed file was converted to a wig file, for visualisation in the Integrative Genomics Viewer, IGV using the following command: bedtools genomecov –bga –ibam brief_name_coverage.bed > brief_name_coverage.wig. However, for Pu-seq data the sequencing reads were further selected post-mapping by samtools Flag-values, as described above. Therefore, to examine genome coverage of those reads used for final Pu-seq analysis an alternative command was used on the ribonucleotide files in the format output_polymerase_strand_correct.bg: awk 'BEGIN{OFS="\t"}{if (NR >=2) print $1,$2,$3,$4}' 6ori_d_r_correct.bg > 6ori_d_r_correct.wig. This created polymerase and strand specific summaries of the number of mapped 5’ ends in the raw (non-normalised) sequencing data as wig files which were also visualised in IGV.

3 Investigating genetic interactions between *RTS1* and replication initiation factors

The findings presented in this chapter have been published (Appendix 8.2) (Wallis & Nieduszynski 2018).

3.1 Introduction

Post-licensing, two kinases, Dbf4-dependent kinase (DDK) and cyclin-dependent kinase (CDK), convert MCM double hexamers to active helicases via the recruitment of firing factors (including Cdc45, GINS and Dpb11). DDK substrates include Mcm4 (Sheu & Stillman 2010); CDK substrates include Sld3 and Sld2 (Zegerman 2015). However, this kinase-driven view of replication initiation is now known to be incomplete, as a role for Rif1-PP1-dependent dephosphorylation in controlling DNA replication initiation has been established (Figure 3.1) (Davé *et al.* 2014; Hiraga *et al.* 2014; Mattarocci *et al.* 2014; Poh *et al.* 2014).

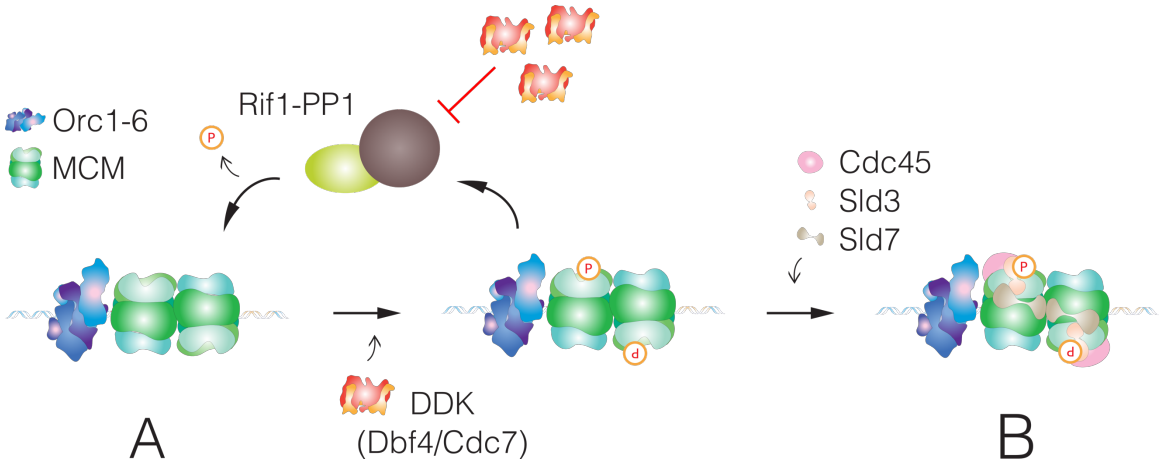


Figure 3.1 Rif1-PP1-opposed DDK phosphorylation of MCM. *(A) Low-level MCM phosphorylation due to DDK activity in G1 or early S phase is removed by Rif1-PP1, this keeps origins in a licensed state, unable to fire. (B) During S phase higher DDK levels lead to MCM phosphorylation and Rif1-PP1 inhibition. This acts as a negative feedback loop giving a switch-like rapid increase in origin firing factor recruitment and origin firing.*

Temperature-sensitive (ts-) replication initiation factor alleles of genes including *DPB11*, *CDC45* and *SLD3* have been observed to have partially rescued phenotypes when *RIF1* was also deleted (Mattarocci *et al.* 2014). A restoration of cell viability due to further genetic modification is a positive (genetic) interaction (Figure 3.2), which results from antagonistic interactions between gene products (in this case between Rif1-PP1 and DDK phosphorylation of MCM).

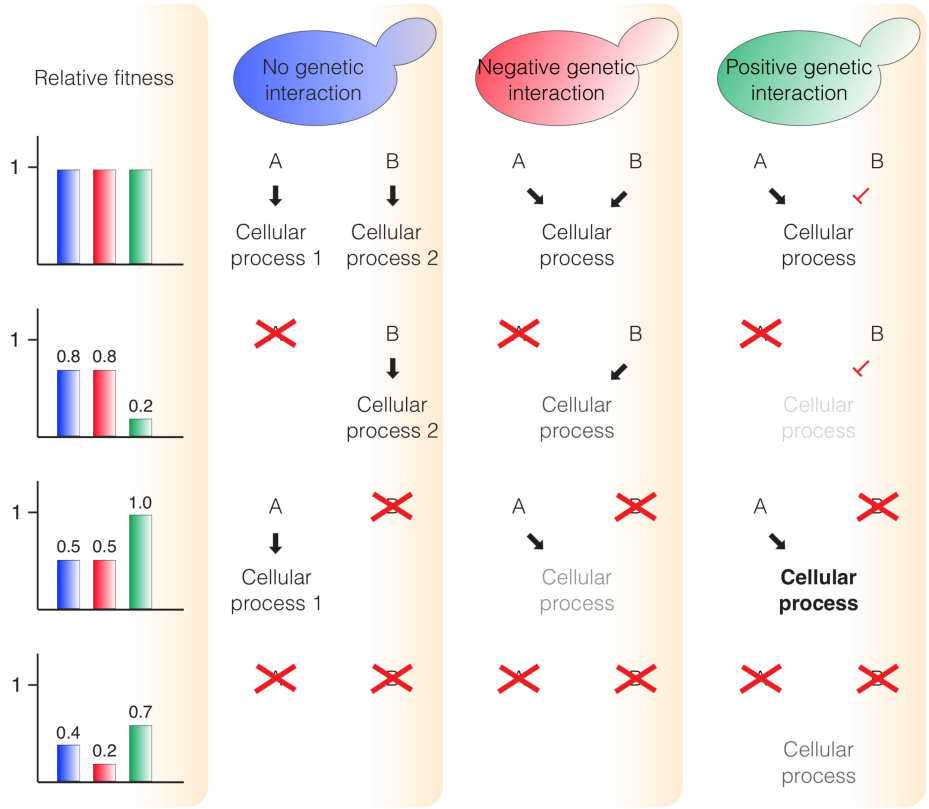


Figure 3.2 A summary of genetic interactions. *Bar charts show the relative fitness in the situations of no genetic interaction (blue), a negative genetic interaction (red) and a positive genetic interaction (green) for; both A and B functioning, A not functioning, B not functioning and A and B not functioning (top to bottom). When A and B function in different cellular processes, the relative fitness of a double mutant is a product of the relative fitness of the two single mutants (no genetic interaction). If the double mutant has a lower than expected viability, it is a negative genetic interaction, indicating that the two proteins function in the same cellular process. In contrast, a greater than expected viability (a positive genetic interaction) indicates that the proteins have opposing roles in a cellular process. In either of the single mutant strains, the cellular process is dramatically altered. However, when both A (stimulation) and B (inhibition) are removed, the resulting double mutant phenotype is less pronounced.*

A genome-wide screen for genetic interactions previously identified a potential positive interaction between mutations in *ORC2* and *RTS1* (Costanzo *et al.* 2010). Rts1 is a regulatory subunit of the PP2A phosphatase (Figure 3.3), which has been previously implicated in metazoan DNA replication control (Chou *et al*. 2002) (Yan *et al.* 2000). This chapter therefore investigated a possible role for Rts1-PP2A in the regulation of DNA replication licensing and firing in *S. cerevisiae*.

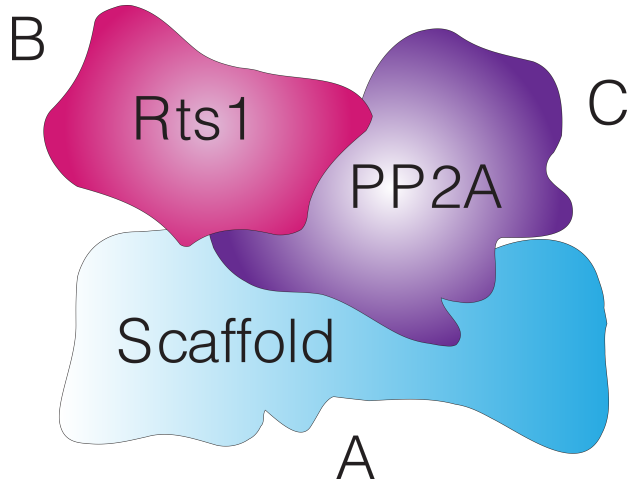


Figure 3.3 A schematic of the PP2A phosphatase structure. *PP2A has a trimeric structure, featuring the catalytic ('C') subunit, a regulatory ('B') subunit, of which there are only 13 alternative subunits in humans, and three in yeast (Including Rts1), and an additional scaffold ('A') subunit* (Stark *et al.* 2015).

3.2 Generating Δ*rts1* ts-origin factor double mutant strains

Many genes encoding DNA replication factors are essential. In order to study DNA replication, temperature-sensitive (ts-) mutants of origin licensing and firing

factors were therefore used. Ts- replication initiation factors are able to function at their permissive temperature, but not at their restrictive temperature when the protein becomes unstable. A panel of ts-mutant yeast strains was assembled to screen for genetic interactions between essential replication initiation factors and *RTS1*. The permissive and restrictive temperatures for each strain were established (Table 3.1); alleles of *CDC45* are shown in Figure 3.4 as an example.

Table 3.1 Temperature-sensitive strains used in this study. *Mutant forms of replication initiation factors, temperatures at which permissive or restrictive phenotypes, and the study that originally reported each strain. The allele cdc45-1 is cold-sensitive, whereas all other alleles are temperature-sensitive.*

REPLICATION FACTOR ALLELE	*OBSERVED RESTRICTIVE TEMPERATURE*	*OBSERVED PERMISSIVE TEMPERATURE*	*REFERENCE*
orc2-1	30	23	Foss *et al*. 1993
cdc6-1	33	30	Hartwell *et al*. 1973
cdc7-4	30	23	Hartwell *et al*. 1973
dbf4-1	32	30	Johnston & Thomas 1982
cdc45-27	32	30	Kamimura *et al*. 2001
cdc45-1	23	30	Moir *et al*. 1982
dpb11-24	37	32	Tanaka *et al*. 2007
cdc9-1	30	23	Hartwell *et al*. 1973

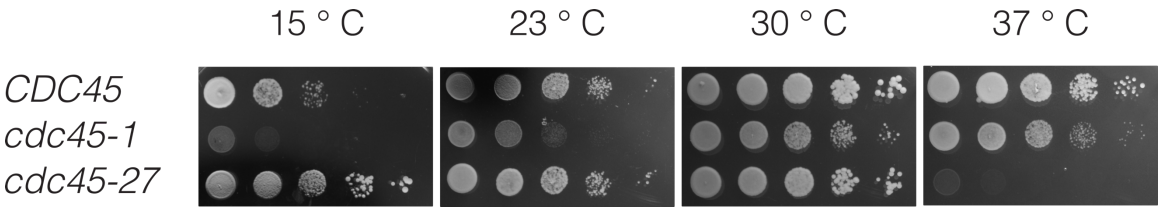


Figure 3.4 Temperature-sensitive* CDC45 *strains grown at permissive and restrictive temperatures. *A comparison of growth at different temperatures of budding yeast strains which are* CDC45 *(wild-type),* cdc45-1 *(cold-sensitive) or* cdc45-27 *(heat-sensitive).*

In order to examine genetic interactions of ts-replication initiation factors and *RTS1*, the *RTS1* gene had to be inactivated. This was achieved by amplifying a *KanMX* cassette from the *S. cerevisiae* Genome Deletion Project (SGDP) and transforming it into wild-type strain. SGDP strains each have one non-essential open reading frame replaced with a *KanMX* cassette. Since the SGDP strains have an S288c background, the *KanMX* cassette which deleted *RTS1* in the SGDP was PCR amplified and transformed into a W303 wild-type strain (Figure 3.5). This strain was subsequently mated to the ts-replication initiation factor strains and haploid double mutants from meiotic spores, which were both temperature-sensitive and G418 resistant were selected (Figure 3.6). An identical approach was used to create a parallel panel of ts-replication initiation factor strains with *RIF1* deleted, which served as a positive control.

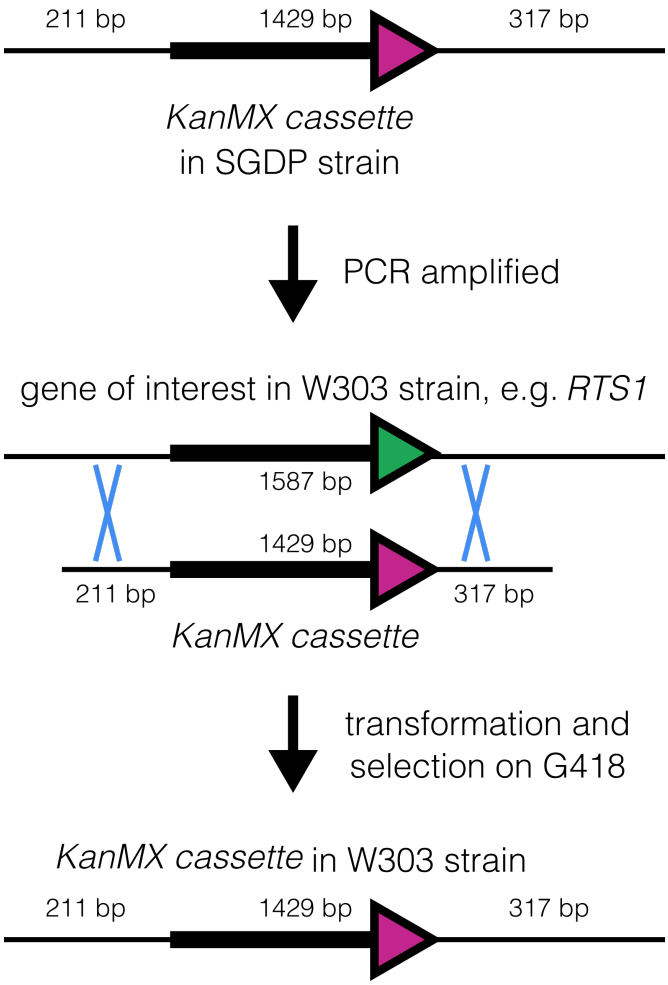


Figure 3.5 Schematic of the approach used to delete open reading frames using* KanMX *cassettes amplified from SGDP strains. KanMX *cassettes from SGDP strains were PCR amplified and transformed into wild-type W303 strains. Transformants were selected on G418 plates and confirmed by appropriate diagnostic PCRs.*

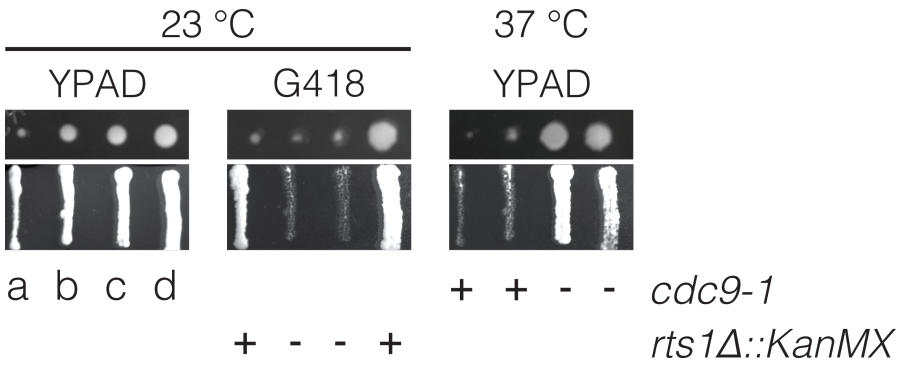


Figure 3.6 An example of the approach used to select double mutant strains. *After tetrad dissection plates were grown at 23 °C, which resulted in four sister colonies (a-d). After replica plating, growth on G418 determined the strains carrying a* KanMX *cassette (*∆rts1::KanMX*). After replica plating, inability to grow at 37 °C indicated those strains with a ts-mutation (*cdc9-1*). Colony 'a' is the double mutant in this example.*

3.3 *RTS1* deletion does not rescue *orc2-1* temperature-sensitivity

The combination of two mutations, which individually decrease cell fitness, can restore fitness if the genes have opposing effects (a positive genetic interaction, Figure 3.2). In the presence of an origin licensing factor mutant allele, such as *orc2-1*, there is a growth deficiency due to reduced origin licensing and subsequent firing. Even at the permissive temperature for *orc2-1*, origin firing can be reduced by as much as 30% (Shimada *et al.* 2002). *RTS1* deletion was first examined in an *orc2-1* strain (Figure 3.7A), since a positive interaction had been reported (Costanzo *et al.* 2010). A dilution series viability assay was used to assess genetic interactions. It was found that an *orc2-1* ∆*rts1* strain had a slightly

increased ts-phenotype relative to *orc2-1* or ∆*rts1* single mutant strains, consistent with no genetic interaction (Figure 3.7A, compare rows 3 and 4). *RIF1* deletion is known to partially rescue ts-phenotypes of some replication initiation factor mutants, including *orc5-1* (Mattarocci *et al.* 2014). Consistent with these previous observations, a small reduction in temperature-sensitivity in an *orc2-1* ∆*rif1* strain relative to an *orc2-1* strain was observed (Figure 3.7A, compare rows 4 and 5).

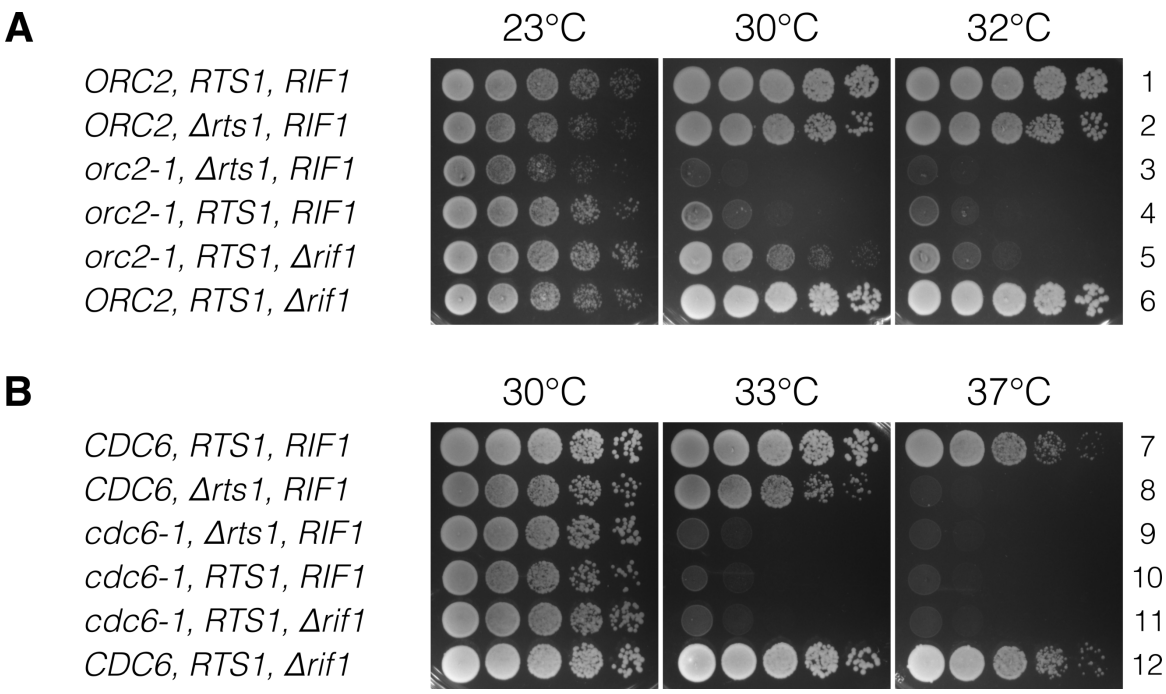


Figure 3.7* RTS1 *deletion does not suppress temperature-sensitivity of DNA replication origin licensing factor mutants. *Budding yeast strains with ts-replication initiation factors, together with either* ∆rts1 *or* ∆rif1 *were characterised by dilution viability assays. Wild-type strains, without ts-replication initiation factors, are shown at the top of each panel, as a control. (A) An ORC mutant allele (*orc2-1*) was assayed. (B) A second pre-RC mutant allele (*cdc6-1*) was assayed.*

3.4 *CDC6* does not interact genetically with *RTS1*

RIF1 deletion provides limited or no rescue to the temperature-sensitivity of pre-RC factor mutants, since they are defective in functions prior to DDK phosphorylation – the step opposed by Rif1-PP1 (Mattarocci *et al.* 2014). A putative role for Rts1 in origin firing could occur either before or after the DDK step. Therefore, a potential genetic interaction between *RTS1* and another pre-RC factor, *CDC6*, was assessed. Deleting *RTS1* in the context of *cdc6*-1 showed no rescue of temperature-sensitivity relative to the original ts-strain (Figure 3.7B, compare rows 9 and 10). Similarly, deletion of *RIF1* gave no positive interaction, consistent with published data (Figure 3.7B, compare rows 10 and 11) (Mattarocci *et al.* 2014). It has previously been shown that Δ*rts1* strains are ts- at 37 °C (Auesukaree *et al.* 2009; Shu & Hallberg 1995), while Δ*rif1* strains are not (Mattarocci *et al.* 2014); the results presented here confirmed both of these phenotypes (Figure 3.7B). A caveat to this finding is that, whilst the *cdc6-1* strain grew at a level similar to wild-type at 30 °C, the strain appeared unable to grow at 33 °C (Figure 3.7, row 10). This means there was no semi-permissive temperature examined to determine if deleting *RTS1* would increase the temperature-sensitivity slightly in a *cdc6-1* strain, as it did in an *orc2-1* strain.

3.5 Δ*rts1* does not antagonise or promote DDK phosphorylation

A potential role for Rts1-PP2A phosphatase in DNA replication initiation could be to oppose the action of a kinase. The established role for Rif1-PP1 in opposing

DDK activity indicates that regulation of phosphorylation is a key step during replication initiation. Therefore, *RTS1* was deleted in combination with ts-alleles of the genes encoding both subunits of DDK (Cdc7 and Dbf4). Deletion of *RTS1* slightly increased the temperature-sensitivity of a *cdc7-4* strain (Figure 3.8A, compare rows 3 and 4) and had no effect in a *dbf4-1* strain (Figure 3.8B, compare rows 9 and 10). This was in contrast to the restoration of cell growth at elevated temperatures upon deletion of *RIF1* in these strains (Figure 3.8A and B, compare

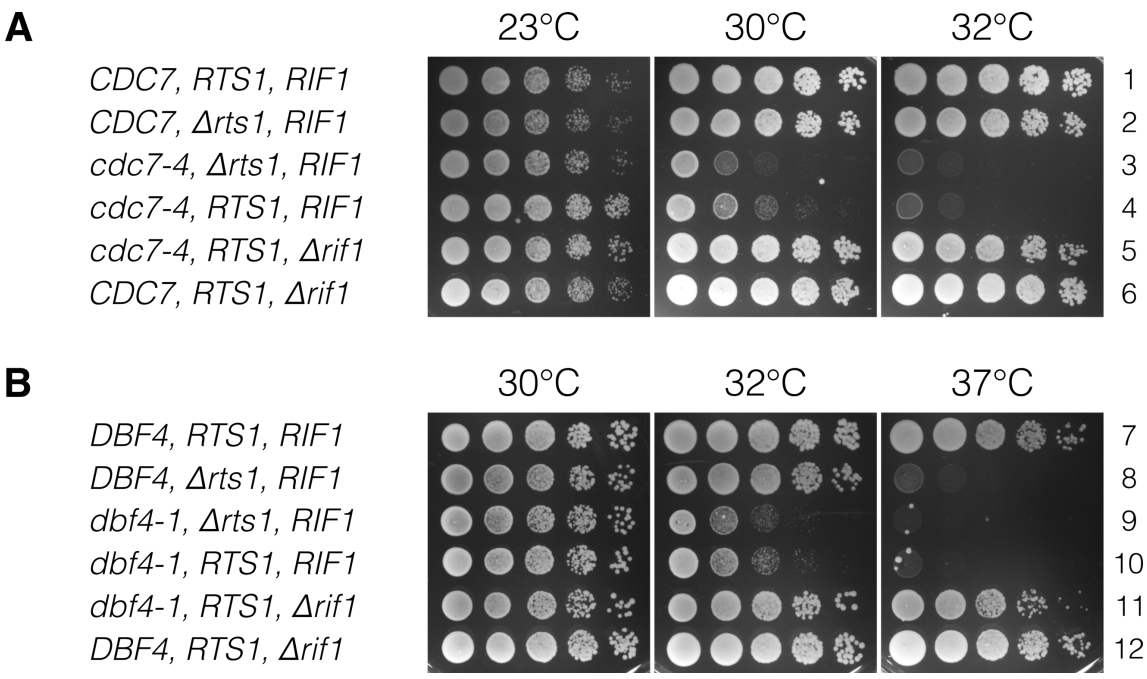


Figure 3.8 Unlike* Δrif1*,* Δrts1 *cannot synthetically rescue ts-forms of DDK subunits. *Budding yeast strains with ts- replication initiation factors, together with either* Δrts1 *or* Δrif1 *were characterised by dilution viability assays. Wild-type strains, without ts-replication initiation factors, are shown at the top of each panel, as a control. Ts-alleles of DDK subunits,* Cdc7 *(A) and* Dbf4 *(B) were assayed.*

rows 4 and 5, and 10 and 11). This suggests that Rts1-PP2A neither promotes nor antagonises DDK activity, whilst confirming the known antagonistic role of Rif1 (Mattarocci *et al.* 2014).

3.6 Deletion of *RTS1* increases the temperature-sensitivity of replication firing factor mutants

MCM phosphorylation by DDK recruits the firing factor Cdc45. The sequential recruitment of further firing factors (e.g. Dpb11) relies upon CDK phosphorylation activity. Therefore, Rts1-PP2A could be important in steps subsequent to the action of DDK. In order to test this hypothesis, *RTS1* and *RIF1* were deleted in the context of ts- *cdc45-1, cdc45-27* and *dpb11-24* firing factor mutant alleles (Figures 3.9 A, B and C, respectively). As previously reported, Δ*rif1* slightly rescued *cdc45-27* and *dpb11-24* temperature-sensitivity, indicating positive interactions (Figure 3.9A and 3.9C, compare rows 4 and 5, and 16 and 17) (Mattarocci *et al.* 2014). Conversely, Δ*rts1* led to increased temperature-sensitivity in these strains, suggesting negative interactions (Figure 3.9A and 3.9C, compare rows 3 and 4, and 15 and 16).

However, the cold-sensitive *cdc45-1* strain (restricted at 15°C) had no phenotypic rescue by either Δ*rif1* or Δ*rts1* (Figure 3.9B, compare row 10 with rows 9 and 11). Instead, increased temperature-sensitivity was observed for both gene deletions, which was strongest in the case of *RTS1*. This additive effect of Δ*rif1* in the cold-sensitive *cdc45-1* strain contrasts with the rescue of *cdc45-27* by Δ*rif1* (Figure 3.9A) (Mattarocci *et al.* 2014). Therefore, the significance of an additive effect of Δ*rts1* with *cdc45-1* is unclear.

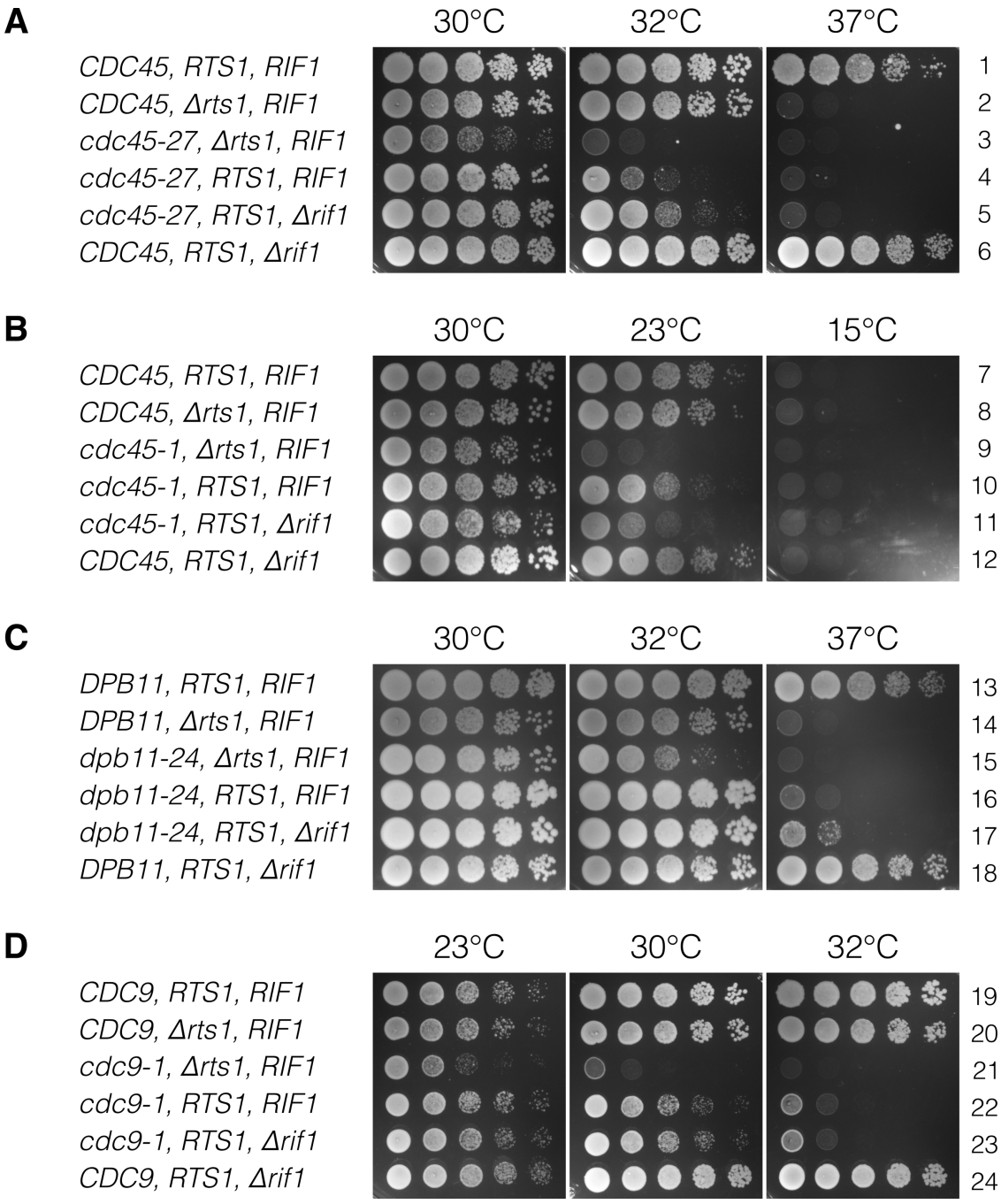


Figure 3.9* RTS1 *has a negative genetic interaction with DNA replication firing factors and a DNA replication progression factor. *Budding yeast strains with ts- replication initiation factors, together with either* Δrts1 *or* Δrif1 *were characterised by dilution viability assays. Wild-type strains, without ts-replication initiation factors, are shown at the top of each panel, as a control. (A) Heat-*

sensitive (I) and cold-sensitive (II) forms of Cdc45, (B) the replication firing factor Dbp11, which functions post-DDK, and (C) Cdc9 ligase, were assayed.

3.7 *RTS1* deletion increases the temperature-sensitivity of a *cdc9* ligase mutant

Since the detrimental effect of *RTS1* deletion alongside ts-replication initiation factors appeared to be most severe in factors which function later in the firing process, it was hypothesised that Rts1 could play a role in DNA replication after origin firing. Therefore, *RTS1* was deleted alongside a ts-*cdc9* ligase allele (Johnston & Nasmyth 1978). Cdc9 ligates Okazaki fragments, aiding in the completion of lagging strand synthesis. *RTS1* deletion in a *cdc9-1* strain resulted in increased temperature-sensitivity relative to the single mutant strains (Figure 3.9D, compare rows 21 and 22). This was the greatest exacerbation of a temperature-sensitive phenotype for a replication factor mutant upon *RTS1* deletion that was observed. In contrast, no effect of Δ*rif1* was seen in *cdc9-1* cells (Figure 3.9D, compare rows 22 and 23).

3.8 *TRP1* deletion has a slight impact on temperature-sensitivity

To confirm that effects observed in post-DDK replication factor mutants were not due to the process of transforming a *KanMX* cassette into *S. cerevisiae*, a non-phosphatase gene was also deleted in the context of these ts-replication initiation factors. *TRP1* encodes a protein which catalyses a step in the tryptophan amino

acid biosynthesis pathway (DeMoss 1965) and can be used as a selective marker in yeast genetic manipulations (Struhl *et al.* 1979).

Deletion of *TRP1* in the *cdc6*-1 mutant allele strain led to a slightly increased temperature-sensitivity (Figure 3.10A, compare rows 3 and 4). This was akin to the slightly increased temperature-sensitivty observed when *RTS1* was deleted in the *cdc6-1* mutant allele strain (Figure 3.7A). When *TRP1* was deleted alongside *cdc7-1 cdc45-27*, *dpb11-24* and *cdc9-1*, a very slight increase in the ts-phenotype was observed (Figure 3.10B, compare rows 7 and 8, 3.10C, compare rows 11 and 12, 3.10D, compare rows 15 and 16, 3.10 E, compare rows 19 and 20). These data indicated that introduction of a *KanMX* cassette alone may slightly increase temperature-sensitivity.

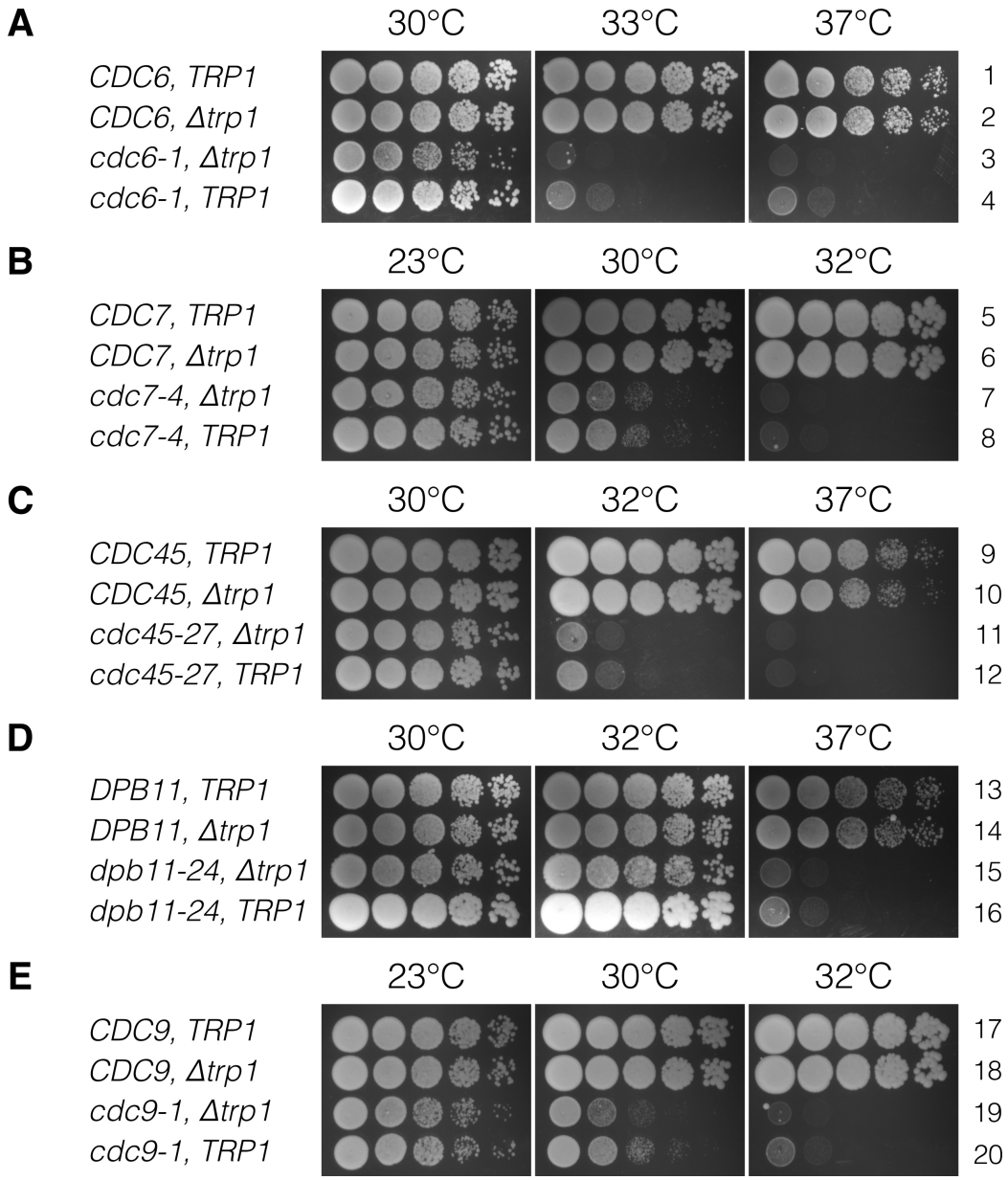


Figure 3.10* TRP1 *has no genetic interaction with DNA replication factors. *Budding yeast strains with ts-replication initiation factors, together with either* TRP1 *or* Δtrp1 *were characterised by dilution viability assays. Wild-type strains, without ts-replication initiation factors, are shown at the top of each panel, as a control.*

3.9 Summary

This thesis aimed to examine the regulation of DNA replication origins and the consequences for DNA replication when origins are de-regulated. This chapter specifically investigated a potential role for Rts1 in regulating the ability of origins to initiate DNA replication. Here, *RTS1*, which encodes a regulatory subunit for the PP2A phosphatase, was deleted in the context of a range of ts-replication factor mutants. As a control, *RIF1* was also deleted. Rif1 recruits PP1 phosphatase to DNA replication origins where it counteracts MCM phosphorylation by DDK. Replication origin firing is reduced in replication factor mutants, however additionally removing negative regulation by Rif1-PP1 has been shown to suppress ts-phenotypes (Mattarocci *et al.* 2014). Deletion of *RIF1* was confirmed to reduce the temperature-sensitivity of mutants in DDK subunits (*cdc7-4* and *dbf4-1*) and some replication firing factors. In contrast, *RTS1* deletion did not rescue temperature-sensitivity in any of the replication factor mutants examined. The lack of positive interaction between *RTS1* and replication initiation factors was consistent with no role for Rts1 in limiting origin firing, analogous to that played by Rif1.

Deletion of *RTS1* resulted in increased temperature-sensitivity of *orc2-1*, *cdc7-4*, *cdc45-27* and *dpb11-24* replication factor mutants. In the absence of a genetic interaction, combining mutations in genes involved in separate pathways will give an additive effect (Figure 3.2). Alone, Δ*rts1* is known to result in lethality above 34°C and a mild reduction in growth at lower temperatures. Therefore, the slightly increased temperature-sensitivity observed in *orc2-1* Δ*rts1* and *cdc7-4* Δ*rts1* strains was consistent with an additive effect and no genetic interaction (Figure 3.7A, 3.8A). Additionally, the increased temperature-sensitivity observed

in the *cdc7*-4, ∆*rts1* was of a similar extent to the increased temperature-sensitivity seen by deletion of *TRP1* in the *cdc7-4* mutant allele strain (Figure 3.10B). *KanMX* insertion into the *TRP1* locus was used as a negative control for *KanMX* insertion, since Trp1 is not involved in DNA replication initiation. Additionally, the *TRP1* locus in W303 yeast contains a non-functional allele (*trp1-1*) and therefore the deletion of this with *KanMX* does not produce a phenotype (Matheson *et al.* 2017).

However, the extent of reduced cell viability seen in *dpb11-24* ∆*rts1* and *cdc45-27* ∆*rts1* double mutant strains (Figure 3.9A, 3.9C), suggested a negative interaction. This may be the result of non-specific protein instability after heat stress in ∆*rts1* strains. Over-expression of *RTS1* can partially rescue lethality of a ts-*HSP60* allele (Shu and Hallberg, 1995). Hsp60 is a mitochondrial protein that aids in protein refolding after heat stress (Shu and Hallberg, 1995). A reduced capability to maintain protein structure under heat stress conditions could explain the increased temperature-sensitivity of unstable replication factors in ∆*rts1* cells. This could be analogous to the partial rescue of the ts-phenotype of *orc2-1* by mutations in the ubiquitin ligase *UBA1* (Shimada *et al.* 2002). Therefore, the variability in increased temperature-sensitivity upon *RTS1* deletion could be a consequence of variable heat-instability of the replication factor proteins resulting from mutant alleles. Investigation of heat stress in ∆*rts1* strains would be needed to test this hypothesis and elucidate molecular mechanisms. Alternatively, these data may represent a genuine negative interaction in the *dpb11-24* ∆*rts1* and *cdc45-27* ∆*rts1* strains, that could indicate a late role for Rts1 in DNA replication initiation. This would be consistent with previous reports of a role for PP2A in loading Cdc45 onto chromatin, in *Xenopus* cell extracts (Chou *et al.* 2002).

Interestingly, the greatest negative interaction was seen between *RTS1* and the replication factor *CDC9* (Figure 3.9D), which encodes DNA ligase. When *cdc9-1* was combined with a *TRP1* gene deletion a slight negative interaction was observed (Figure 3.10), however, the ts-phenotype of the *cdc9-1* ∆*rts1* strain was much more pronounced, suggesting this effect is specific to the function of *RTS1*. Cdc9 acts subsequent to replication initiation and accordingly no positive interaction with *RIF1* was found (Figure 3.9D). The observed negative interaction between *RTS1* and *CDC9* may provide evidence for a role for Rts1 in allowing replication fork progression. In *cdc9-1* strains, increased DNA damage is seen due to the collapse of replication forks. This results in DSBs, and the DNA damage response (DDR) being activated. One pathway for DSB repair is break-induced replication (BIR), which has been found to be activated in *cdc9-1* cells (Vasianovich *et al.* 2014). Therefore, hypothetically, there could be a role for Rts1 in recruiting PP2A phosphatase during BIR. In this case, deleting *RTS1* in a *cdc9-1* background would give increased temperature-sensitivity, since there would be an impaired capacity for cells to repair *cdc9-1* dependent DSBs. Over-expression of *RTS1* could potentially compensate for the increased replication fork collapse seen in *cdc9-1* mutants, giving reduced temperature-sensitivity. Further studies would be needed to test this hypothesis.

Although Rts1 is considered a PP2A subunit, there is some evidence that Rts1 can interact with other phosphatases, including PP1, under glucose-induction (Castermans *et al.* 2012). Therefore, this study cannot rule out the possibility that the genetic interactions considered here refer to phosphatases other than PP2A. However, since the association of Rts1 with other phosphatases

has only been demonstrated in one induced condition, it is considered unlikely that other phosphatases are involved.

Evidence in support of the negative interactions between *RTS1* and DNA replication factors can be found on BioGrid[3.5], a summary of published genetic interactions in budding yeast. A combination of high and low throughput genetic interaction screens published after the work in this chapter was completed showed that *RTS1* exhibits negative interactions with *DBF4*, *CDC6*, *ORC6* and *DPB11* (Archambault *et al.* 2005; Collins *et al.* 2007; Costanzo *et al.* 2016). However, this study has shown, for the first time, negative interactions of *RTS1* with *CDC45* and *CDC9*.

4 Characterising cell growth and DNA replication of *Saccharomyces cerevisiae* with a large replicon

4.1 Introduction

The second part of this thesis examines the consequences of inactivated origins in *S. cerevisiae*. Budding yeast strains which were disomic for chromosome III (ChrIII) or a fragment of ChrIII were used. It has been known for over ten years that despite the absence of active origins a second copy of ChrIII is able to be replicated and maintained in *S. cerevisiae* (4.4×10^{-4} loss/division) at a similar level to a second copy of ChrIII with active origins (3.7×10^{-5} loss/division) (Dershowitz *et al.* 2007). A ChrIII fragment of 174 kb which lacked origins distal to *ARS310* was maintained slightly less efficiently when active origins *ARS305*, *ARS306*, *ARS307*, *ARS309* and *ARS310* were inactivated (2.4×10^{-3} loss/division relative to 7×10^{-5} loss/division, Figure 4.1), showing that dormant origins on the right arm of ChrIII contributed to this maintenance. However, a smaller 142 kb fragment of ChrIII both lacking the active origins listed above and lacking dormant origins *ARS300*, *ARS301*, *ARS302/303* and *ARS304* was also maintained in more than 97 % of cells, even in the absence of the centromeric origin *ARS308* (Dershowitz *et al.* 2007). This suggested an alternative form of replication was occurring. This replication mechanism was not dependent on homologous recombination; ten genes in the *RAD52* epistasis group were found to have no impact on loss rate of the 174 kb fragment lacking active origins (Theis *et al.* 2010).

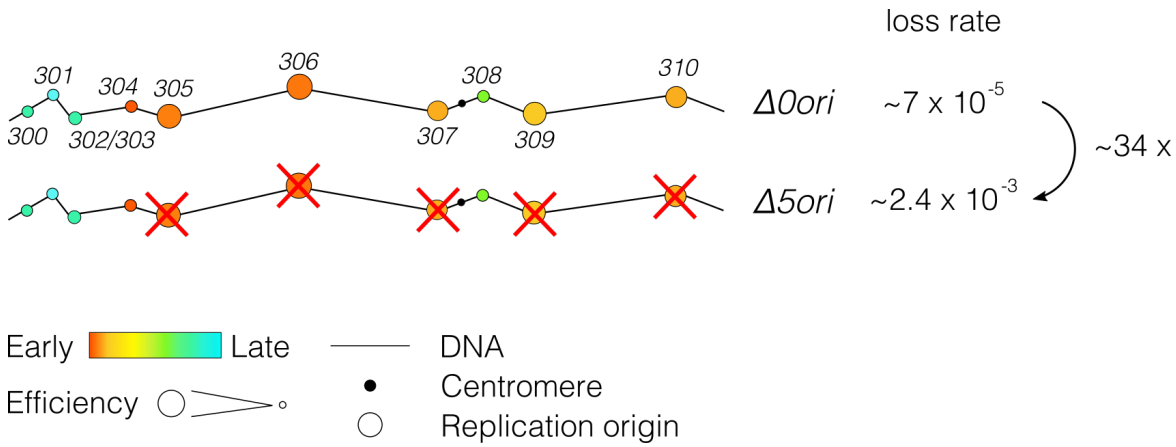


***Figure 4.1 Loss rates of* Δ0ori *and* Δ5ori *174 kb chromosome III fragments from* S. cerevisiae.** *Replication origins are represented by circles and labelled with their ARS number (e.g. 301 =* ARS301*). Circle diameter represents efficiency (the mean of origin efficiencies calculated from replication timing data (Hawkins et al. 2013b) and okazaki fragment data (Smith & Whitehouse 2012) and the colour indicates mean replication time from early (red) to late (blue) based on replication timing data (Hawkins et al. 2013b). Loss rates taken from published data (Dershowitz et al. 2007).*

Following this work, in 2014 it was shown that removing multiple adjacent active origins from a genomic locus was also tolerated by *S. cerevisiae*. The deletion of seven active origins from chromosome VI (ChrVI) did not impair yeast growth rates, despite a reduction in Orc2 binding at the deleted origins (Figure 4.2) (Bogenschutz *et al.* 2014). Whilst Orc2 binding was reduced at the deleted origins, ORC binding and initiation of DNA replication (measured by BrdU incorporation) seemed to occur non-specifically at sites near to deleted origins. However, dormant origins *ARS602.5* and *ARS603.1* significantly contributed to replication in the absence of surrounding origins (Bogenschutz *et al.* 2014).

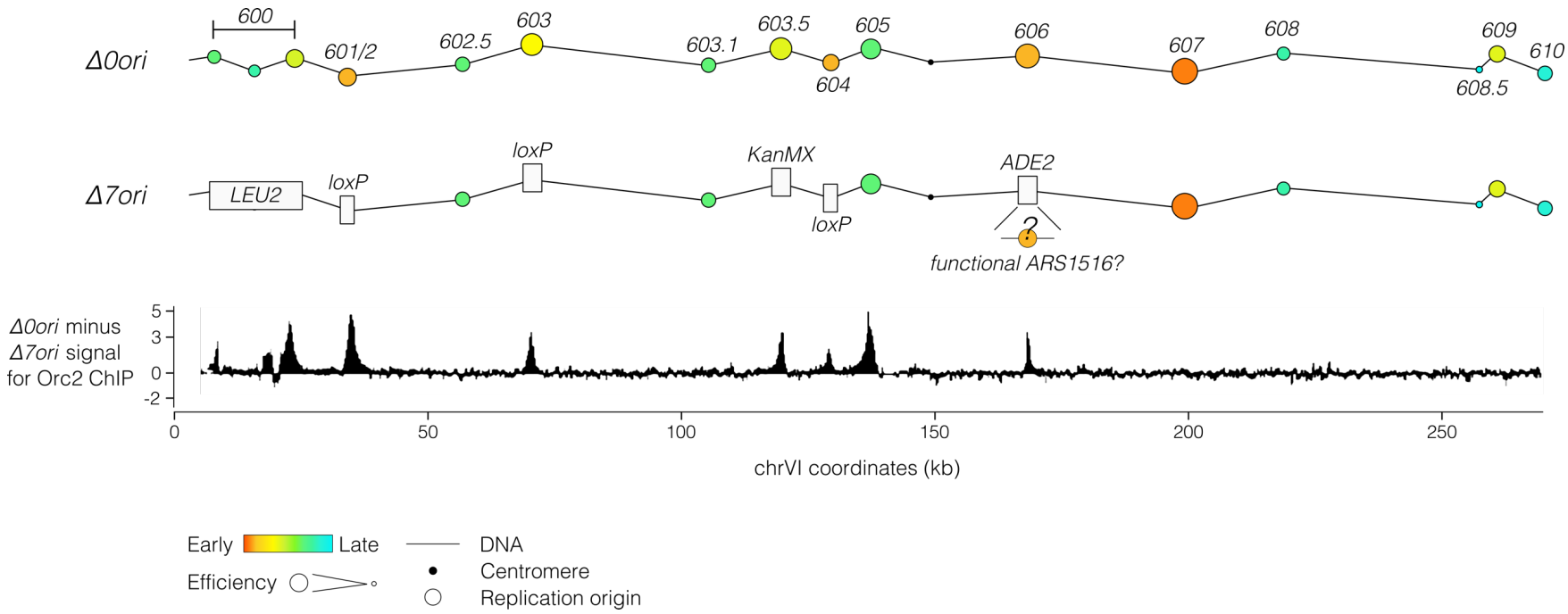


Figure 4.2 Removing seven active chromosome VI origins in* S. cerevisiae *leads to reduced Orc2 binding. *Replication origins are represented by circles and labelled with their ARS number (e.g. 601* = ARS601*). Circle diameter represents efficiency (the mean of origin efficiencies calculated from replication timing data* (Hawkins *et al.* 2013b) *and okazaki fragment data (Smith & Whitehouse 2012)) and the colour indicates mean replication time from early (red) to late (blue) based on replication timing data (Hawkins et al. 2013b). The ADE2 marker contains ARS1516 sequence, although the predicted ACS was not included. Chromosome VI with seven active origins replaced by active markers has reduced capacity to bind ORC, demonstrated by ChIP of Orc2. Image modified from (Bogenschutz et al. 2014).*

A caveat to this finding, was that DNA initiation close to the *ars606* (replaced by *ADE2*) was observed. The *ADE2* gene overlaps with the active origin *ARS1516*. Although the predicted ACS site was not included in the *ADE2* insertion at *ARS606*, the surrounding sequence may have been capable of initiating replication (Bogenschutz *et al.* 2014).

The aim of the work in this chapter is to examine the impact of a large replicon on the growth and DNA replication of *S. cerevisiae*. The yeast strains with a second copy of a 174 kb fragment of ChrIII CB08D3 (Δ*0ori*) and YDN289 (Δ*5ori*) (Dershowitz *et al.* 2007) were procured and used for initial study, before characterising a new genomic large replicon on ChrVI.

4.2 Characterising a disomic chromosome III strain

4.2.1 Disomic strains with inactive origins on the 174 kb ChrIII fragment grow more slowly

S. cerevisiae strains with a 174 kb fragment of ChrIII were obtained (Dershowitz *et al.* 2007). Each strain contained a balancer chromosome from a related yeast species (*Saccharomyces carlsbergensis*) either with all origins active (Δ*0ori*) or with five active origins deleted (Δ*5ori*, Figure 4.1). The balancer chromosome contained multiple genetic markers including *his4*-S8, which complemented the genomic *his4-280*, meaning cells could produce histidine only when both copies of ChrIII were maintained. Therefore, strains were grown in –HIS medium at all times. A dilution series assay showed that the Δ*5ori* strain had a slight reduction in colony growth relative to the Δ*0ori* strain at 37 °C (Figure 4.3). A slight reduction in growth was also observed at 25 °C and 30 °C (Figure 4.3). This indicated there

was no significant difference in cell viability between the ∆*5ori* strain and the ∆*0ori* strain.

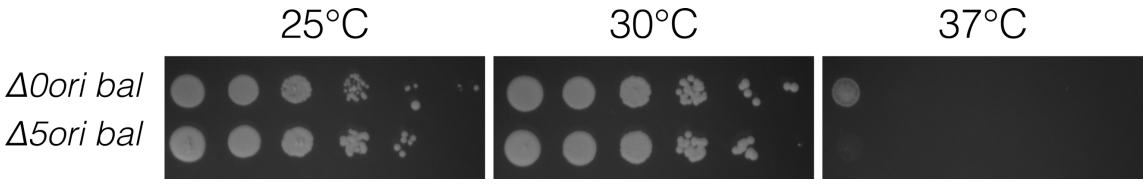


Figure 4.3 Colony growth at a range of temperatures in the* ∆5ori *strain. *A series dilution assay was carried out with the most concentrated cells at 10^7 cells/ml and serial ten-fold dilutions. Cells were grown on -HIS plates at stated temperatures for two days prior to imaging.*

However, when cells were grown in liquid cultures, the mean generation time at 30 °C of the ∆*5ori* strain (164 min) was significantly longer than that of the ∆*0ori* strain (133 min) (student's t-test, two-tailed, $p < 0.01$, Figure 4.4, 4.5).

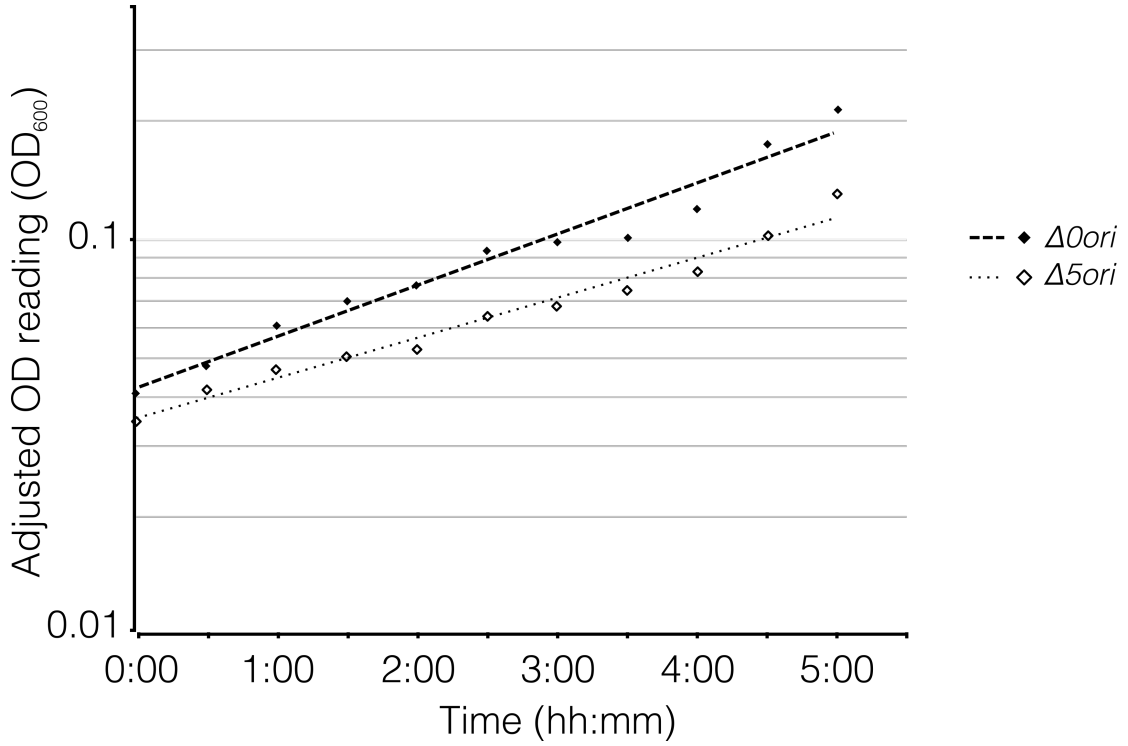


Figure 4.4 Example growth curves for yeast strains with* ∆0ori *and* ∆5ori *ChrIII 174 kb fragments. *OD_{600} readings from yeast grown in liquid –HIS medium at 30 °C were taken over five hours and plotted against time. The gradient of the line of best fit (see key) indicates rate of growth.*

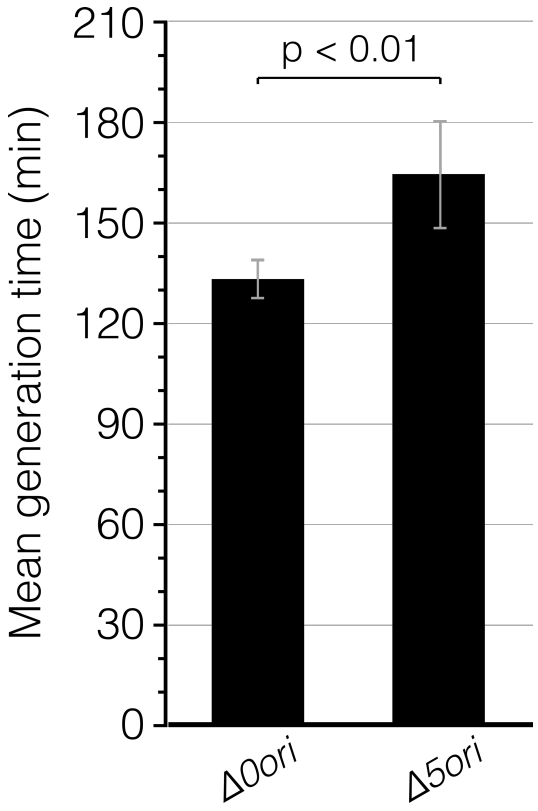


Figure 4.5 The mean generation time of* S. cerevisiae *strains with a chromosome III 174 kb fragment is increased when five origins are inactivated. *Mean replication time was calculated from growth curve experiments in liquid -HIS medium at 30 °C. Five repeats per strain were carried out, and error bars represent one standard deviation. A significant difference was found between the two mean generation times (Student's t-test, two-tailed, p < 0.01).*

4.2.2 DNA replication of a disomic chromosome III strain

Preliminary work by a lab member, Carolin Müller, had investigated DNA replication in the ChrIII *Δ5ori* 174 kb fragment strain. A synchronous cell cycle experiment was performed, where exponentially growing cells were arrested in G1 then synchronously released into S phase. This showed that the DNA in most

of the genome was replicated according to the same temporal programme as wild-type *S. cerevisiae* W303 cells. Equivalent time-points were calculated based on total percentage of DNA replicated to allow direct comparison between strains. At the equivalent of 50 minutes replicated in the W303 strain, the 174 kb ChrIII fragment in the Δ*5ori* strain had reached a relative copy number of approximately 1.4 (Figure 4.6). This was in contrast to the W303 ChrIII, which was at a copy

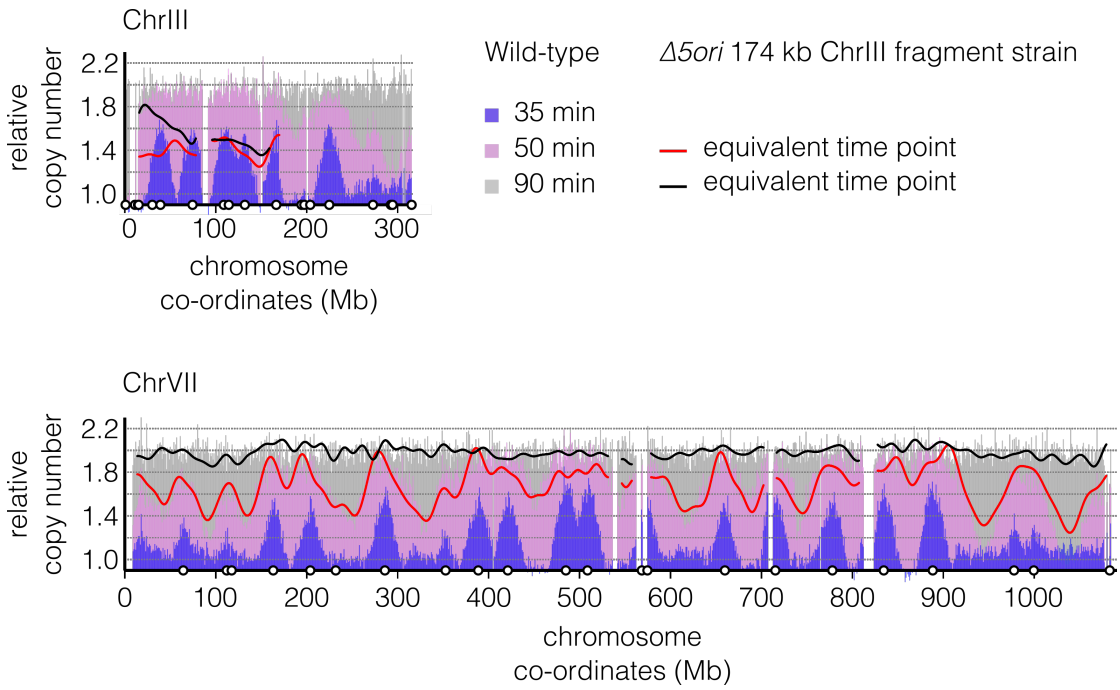


Figure 4.6 ChrIII fragments with inactivated origins replicate later in S phase than wild-type ChrIII. *A synchronous culture of wild-type W303* S. cerevisiae *was sampled at three time points as shown and DNA replication was monitored by next-generation sequencing (see key). The* Δ5ori *ChrIII fragment strain was analysed in the same way and equivalent time points selected using the percentage of DNA replicated. Chromosome VII is shown as a control. Figure (modified) and data courtesy of Carolin Müller (Nieduszynski lab), unpublished data.*

number of approximately 2.0 across the 174-kb region of interest. This indicated a delay of replication on ChrIII in the absence of active origins. No clear peaks were seen which indicated there were no discrete sites of DNA replication initiation at this time point.

After 90 minutes the entirety of W303 ChrIII and chromosome VII (ChrVII) were at a copy number of 2.0, as was ChrVII in the ∆*5ori* strain. However, the ChrIII fragment in the ∆*5ori* strain had a peak relative copy number of 1.8. This peak in replication occurred towards the far left of chromosome III, where the dormant origins (*ARS300* – *ARS304*) were present (Figure 4.1, 4.6). Therefore, known dormant origins appeared able to initiation replication on a disome of chromosome III late in S phase. However, known dormant origins could not explain the approximately 40 % replication observed at the equivalent 50-minute time point

The use of a disome required the yeast to be kept under selective pressure, which is not a natural state. Additionally, ChrIII is the least stable chromosome in the yeast genome, as the centromeric sequence of ChrIII is subject to a relaxed spindle assembly checkpoint (Kumaran *et al.* 2013). Since the centromere is included in the 174 kb ChrIII fragment, this could impact on the stability of DNA. Loss of ChrIII would confound findings based on quantitative measures of ChrIII copy number to examine DNA replication. Finally, the inclusion of genetic markers is not ideal. For example, the *ADE2* marker used to inactive *ARS606* by Bogenschutz *et al*. potentially contained functional sequence from the origin *ARS1516* (Bogenschutz *et al.* 2014). Therefore, a genomic large replicon containing precise mutations of the ACS sites at origins was a more ideal system.

4.3 Generating a large replicon on chromosome VI

*4.3.1 Generating a four-origin deletion strain (*Δ4ori*)*

In order to examine the effect of a large replicon in a genomic locus a new large replicon was created on a chromosome of *S. cerevisiae*. Chromosome VI (ChrVI) was selected due to its well characterised nature in the yeast origin database, oriDB (Nieduszynski *et al.* 2007; Siow *et al.* 2012). To avoid origin-specific effects which were already documented in (Bogenschutz *et al.* 2014), the right arm of chromosome VI was selected.

Origins in *S. cerevisiae* were first identified using plasmid assays to examine each region of the yeast genome for the potential to replicate (Kearsey 1983; Nieduszynski & Donaldson 2009). Past members of the Nieduszynski lab had validated select active origins, including those on the right arm of ChrVI, by mutating the ARS sequence and repeating these plasmid assays (Müller 2012). Using this knowledge, strains were generated in the Niedusynski lab prior to this study where inactive ARS sequences were used to replace active ARS sequences in the yeast genome. The *ARS607*, *ARS608* and *ARS609* origins were inactivated by sequential selection for incorporation of a *URA3* marker, then selection for loss of *URA3* in favour of incorporation of mutated DNA at the origin locus (Figure 4.7). This results in only 5 – 6 bp changes to the genome at each origin. The *ARS610* origin existed within the core X repetitive DNA element, which is found in the sub-telomeric regions of all chromosomes (Louis 1995). Therefore, to prevent recombination leading to repair of the origin, it was completely replaced by a mutant core X with a mutated ACS linked to the selectable *URA3* maker.

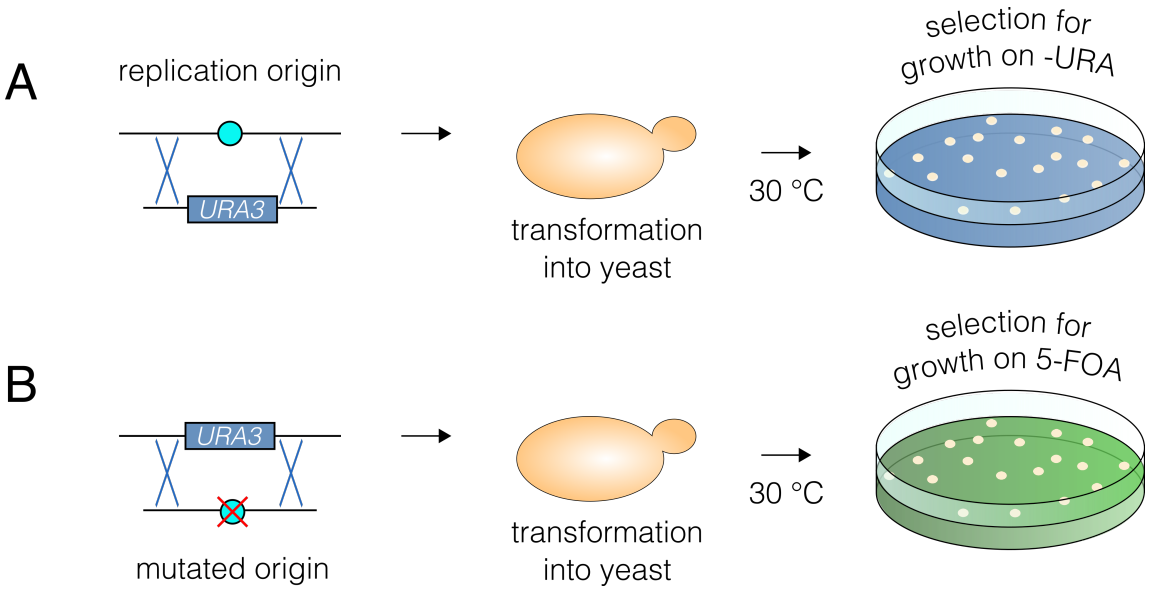


Figure 4.7 Schematic of the approaches used to inactivate yeast origins. *DNA spanning an origin is deleted by homologous recombination with the* URA3 *marker flanked by homologous DNA. This is selected for by growth on –URA medium. The* URA3 *is then deleted in the same manner by mutated origin DNA with homologous sequence and yeast are selected for ability to grow on 5-FOA (5-Fluoroorotic acid) medium.*

This study mated strains containing existing origin inactivations of *ARS607*, *ARS608* (CAY110) and *ARS609*, *ARS610* (AMY337) to create four-origin deletion strains (Δ*4ori*). During the creation of these strains, it was observed that inactivation of either two or four active origins on ChrVI did not affect the growth of sister colonies on YPAD plates (Figure 4.8). Four independent Δ*4ori* strains were generated to confirm there were no colony growth phenotypes in Δ*4ori* strains. Strains of both mating types were generated, accordingly this study proceeded to use ACY162 (*MAT***a**) and ACY165 (*MATα*). It was known that W303

S. cerevisiae strains can contain *rad5* and *bud4* mutations (Elserafy & El-Khamisy 2018; Voth *et al.* 2005). These mutations were screened for in all strains, and only strains that were found to be *RAD5*, *BUD4* were selected for this study. Strain genotypes were confirmed by PCR (spanning and allele-specific) and Sanger sequencing of origins and the surrounding DNA.

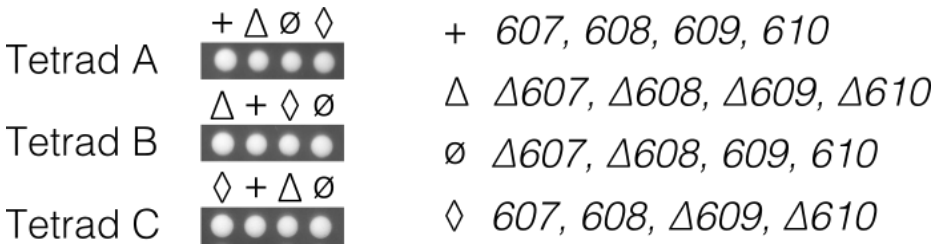


Figure 4.8 Deletion of up to four active ChrVI origins does not affect colony size. *Sister colonies with four, two or no inactivated origins grow to the same size and shape. Three example tetrads are shown. Colonies were grown on YPAD medium for two days prior to imaging.*

4.3.2 Confirming the ACS site of ARS608.5

The ACS sequence of *ARS607*, *ARS608*, *ARS609* and *ARS610* have previously been reported and are listed in oriDB (Nieduszynski *et al.* 2007; Siow *et al.* 2012). The dormant origin, 608.5 was listed in oriDB as a confirmed origin (Figure 4.9), however the functional origin sequence was not known (Nieduszynski *et al.* 2007; Siow *et al.* 2012). *ARS608.5* has a mean replication time of 31.8 minutes (Yabuki *et al.* 2002) which is late compared to early firing origins such as *ARS607* (mean replication time 10 – 20 minutes) (Alvino *et al.* 2007; Raghuraman *et al.* 2001; Yabuki *et al.* 2002). This, combined with its close proximity to *ARS609*

(approximately 3 kb apart) may have contributed to the precise origin site not being previously determined.

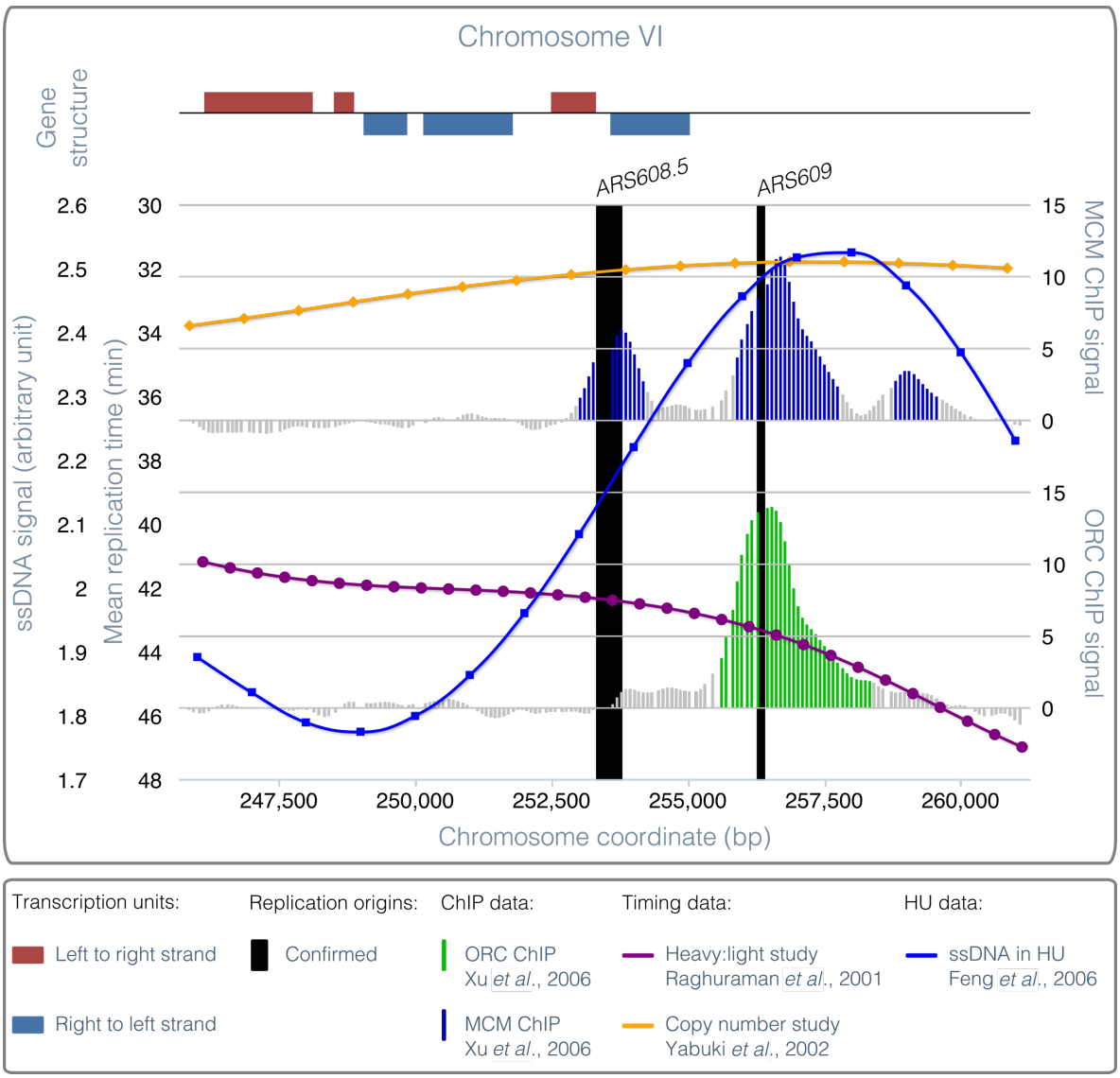


***Figure 4.9 Modified screenshot of oriDB showing published evidence for* ARS608.5.** *oriDB (Nieduszynski* et al. *2007; Siow* et al. *2012) was used to investigate evidence for* ARS608.5 *location and activity. Datasets considered by oriDB are included in the figure legend (Feng* et al. *2006; Raghuraman* et al. *2001; Xu* et al. *2006; Yabuki* et al. *2002).*

During this study, an undergraduate student (Amy Wolstenholme) was supervised for a six-week project investigating dormant origins with unknown functional origin sites, including *ARS608.5*. Briefly, approximately 600 base pairs of DNA surrounding *ARS608.5* was amplified by PCR and ligated into a plasmid vector. Four potential sites which resembled the ACS (Nieduszynski *et al.* 2006) (Figure 4.10) were selected by manual inspection of the DNA in Benchling. For each site, primers overlapping the potential ACS site were designed, which included single base pair mutations designed to inactivate the origin. These primers were used in a PCR with the template *ARS608.5* to create new plasmids, containing mutated potential ACS sites. All four potential sites were mutated individually, and the resulting plasmids used in a preliminary plasmid assay for origin activity. One site prevented colony growth when mutated, whilst three did not alter colony growth (Amy Wolstenholme, personal correspondence; data not shown).

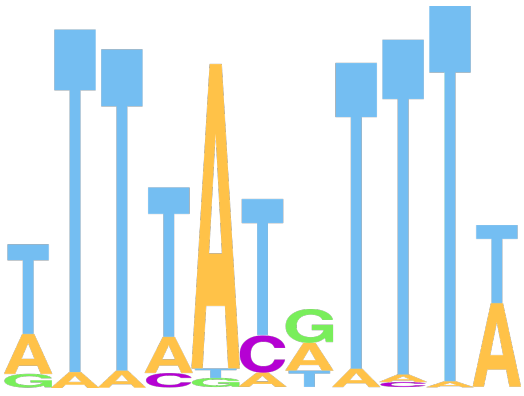


Figure 4.10 ARS consensus sequence (ACS). *Representation of the DNA bases most conserved in the ARS consensus sequence. Modified from (Descorps-Declère et al. 2015).*

4.3.2.1 Verifying a Δ*5ori* strain

The Δ*4ori* strain (ACY162), information on the potential *ARS608.5* ACS site and the mutant *ars608.5* sequence identified by Amy Wolstenholme were made accessible to a collaborator, Vamsi Gali (Donaldson lab), who mutated the *ARS608.5* potential ACS by CRISPR-Cas9 in the Δ*4ori* strain (Vamsi Gali, personal communication). The resulting Δ*5ori* strain (VGY129) Vamsi generated was made available to this study, and this study then carried out a plasmid assay (Figure 4.11) to confirm origin inactivation. The plasmid assay aimed to confirm that the mutated origin sequence in the genome was an inactive origin, with less functional activity than the wild-type *ARS608.5*. The assay was performed using between 600 bp and 1000 bp of DNA amplified from the origin loci and surrounding DNA in either the W303 wild-type or Δ*5ori* strains.

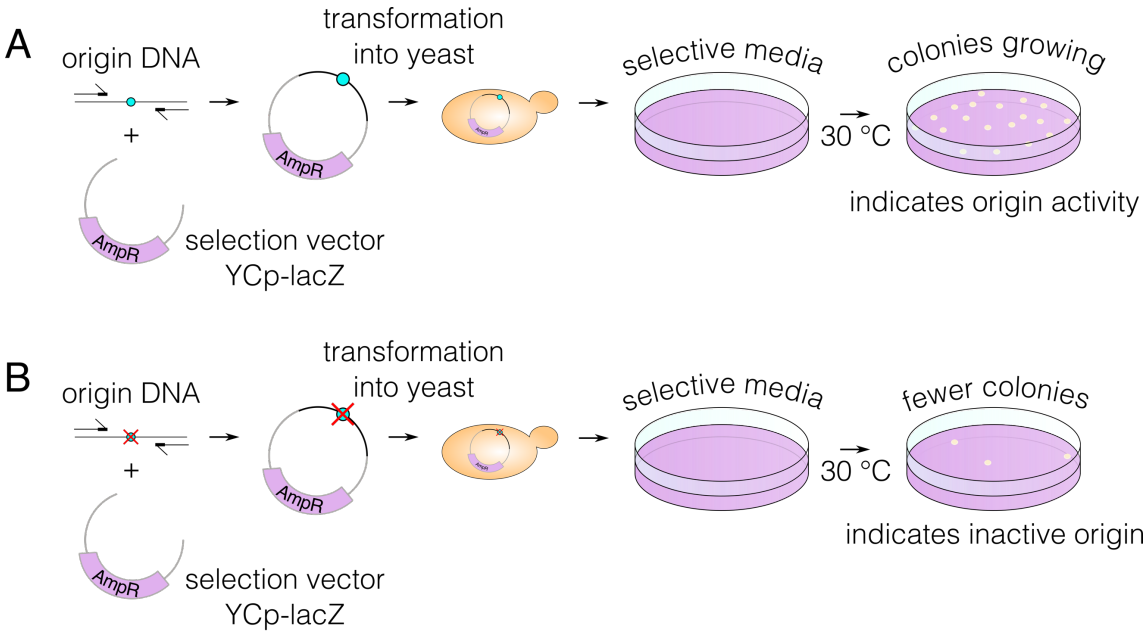


Figure 4.11 Schematic of plasmid assay for determining origin activity. *DNA to be tested for origin activity was ligated into an origin-less selection vector (YCp-lacZ, which contains an genetic marker for ampicillin resistance (AmpR). The new plasmid was transformed into* S. cerevisiae *cells and grown on selective medium for a set number of days. (A) The presence of colonies indicated origin activity whereas (B) no colonies or similar colony growth to the vector only control indicated the DNA had no origin activity.*

ARS608.5 origin activity was weak, only producing colonies which could be seen by eye after incubation for 9 days at 30 °C and two days at room temperature (Figure 4.12). In contrast, the active origin *ARS607* produced large colonies after incubation for 2 days at 30 °C. The existing inactivation *ars607* was used as a control, and as predicted this produced a similar level of colonies to the vector only background level. It was also shown that a mutated form of the potential ACS site identified during this study (*ars608.5*) resulted in reduced

colony growth relative to *ARS608.5*. The level of colony growth was similar to the vector only control and *ars607* (Figure 4.12).

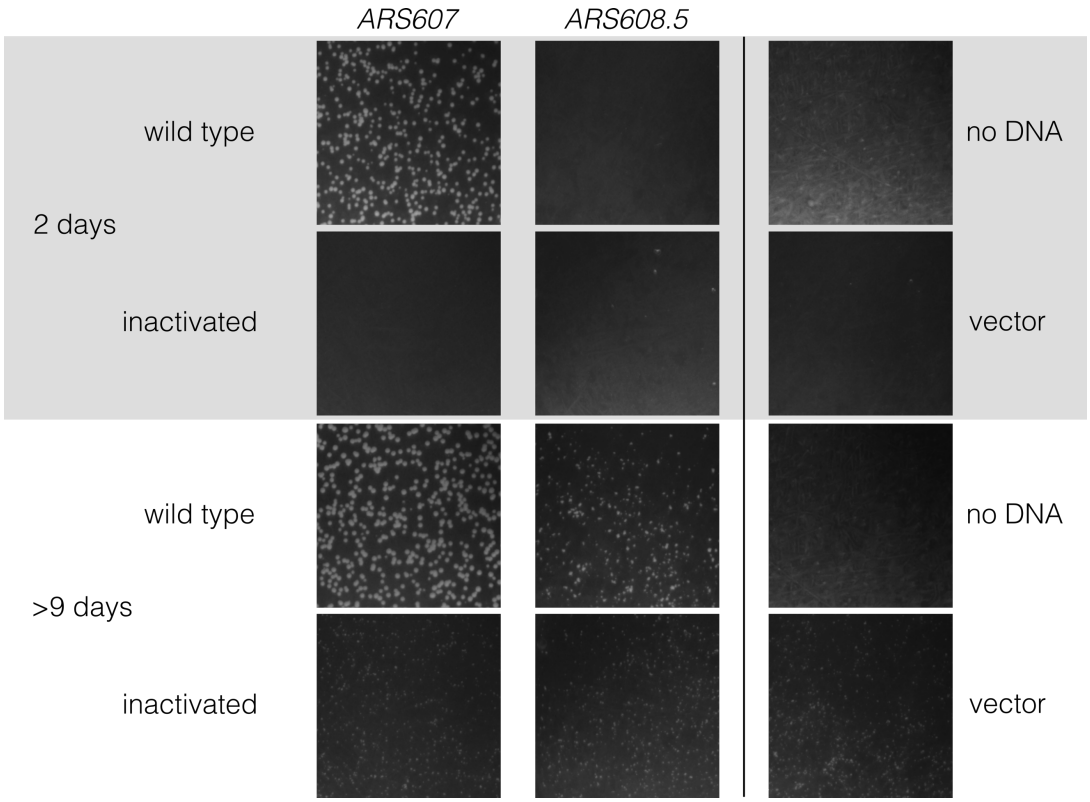


Figure 4.12* ARS608.5 *has weak activity, which is reduced in the inactivated* ars608.5*. *A plasmid assay was carried out, as described in Figure 4.10. Plates were incubated at 30 °C for 9 days, then at room temperature for a further 2 days (where growth was minimal). Plates were imaged after 2 days at 30 °C and after 9 days at 30 °C and 2 days at room temperature. Vector only and no DNA control transformations are also shown.*

4.3.3 Creating Δ6ori *large replicon strains*

The *Δ5ori* strain lacking *ARS607*, *ARS608*, *ARS608.5*, *ARS609* and *ARS610* contained a replicon with no active or dormant origins (as annotated on oriDB) for

over 100 kb (Nieduszynski *et al.* 2007; Siow *et al.* 2012). However, since *ARS606* was known to fire early in S phase, with a mean replication time 12.5 – 20.7 minutes (Alvino *et al.* 2007; Raghuraman *et al.* 2001; Yabuki *et al.* 2002), it was decided to also study a larger replicon where *ARS606* was inactivated.

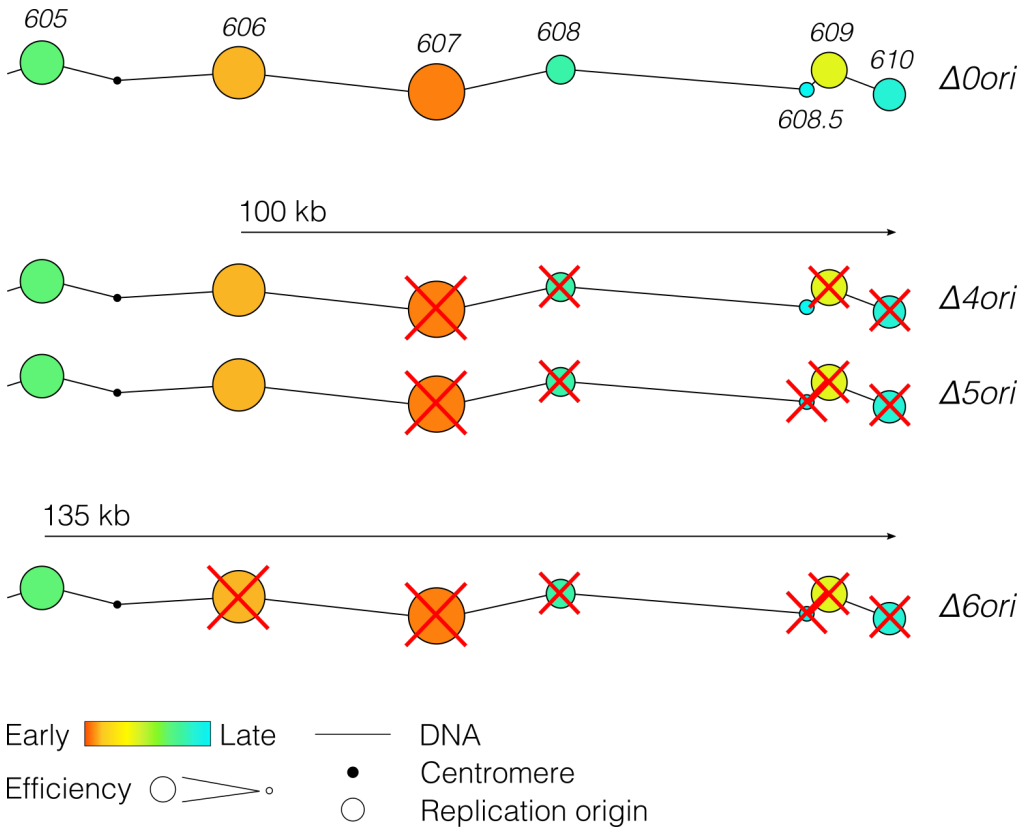


Figure 4.13 Origin deletion strains used in this study. *Replication origins are represented by circles and labelled with their ARS number (e.g. 605 =* ARS605*). Circle diameter represents efficiency (the mean of origin efficiencies calculated from replication timing data (Hawkins et al. 2013b) and Okazaki fragment data (Smith & Whitehouse 2012) and the colour indicates mean replication time from early (red) to late (blue) based on replication timing data (Hawkins et al. 2013b). A section of ChrVI, from* ARS605 *to the telomere is shown.*

The Δ*5ori* strain was therefore mated to a strain lacking *ARS606*, *ARS607* and *ARS608* (CAY109) to produce a Δ*6ori* strain lacking *ARS606*, *ARS607*, *ARS608*, *ARS608.5*, *ARS609* and *ARS610*. All origin and *RAD5* and *BUD4* genotypes were confirmed as in 4.3.1. The additional deletion of *ARS606* extended the replicon almost 35 kb; *ARS605* is located at ~136 kb on ChrVI, the whole chromosome is ~271 kb in length (Figure 4.13). The chromosome VI origin inactivations used during this study are summarised in Table 4.1.

Table 4.1 Origin inactivations used in strains for this study. *Description of origin inactivations present in strains used for this study. Modified bases are underlined.* Xho*I restriction enzyme sites (**CTCGAG**) are in **bold**.*

ORIGIN	*FUNCTIONAL ACS AT THE ORIGIN*	*REPLACEMENT SEQUENCE*	*LOCATION ON CHRVI*
ARS606	AAATATAAAT	A**CTCGAG**AAT	167727-167736
ARS607	TTTATATTTA	**CTCGAG**ATTTA	199417-199426
ARS608	TTTACTTTTAG	TTTAC**CTCGAG**	216485-216495
ARS608.5	TTATACACATAAATA	TTATATA**CTCGAG**TA	253452-253466
ARS609	AAACATAAAA	AA**CTCGAG**AA	256376-256385
ARS610	TTTACGTTTA		269771-269780
		URA3	269729-270161

4.4 A large replicon on ChrVI does not affect cell growth

It had been shown (4.2.1) that inactivation of origins on a second ChrIII fragment led to a reduced growth rate, relative to yeast with active origins on a second ChrIII fragment. However, inactivation of seven origins on the endogenous

chromosome VI was reported to cause no growth defect (Bogenschutz *et al.* 2014). Therefore, the growth of cells with 100 kb and 135 kb replicons on the right arm of chromosome VI was investigated. A dilution series assay showed that the *Δ4ori, Δ5ori* and *Δ6ori* strains had equal colony growth to the *Δ0ori* strain at a range of temperatures (Figure 4.14). This indicated that all four strains had similar viability.

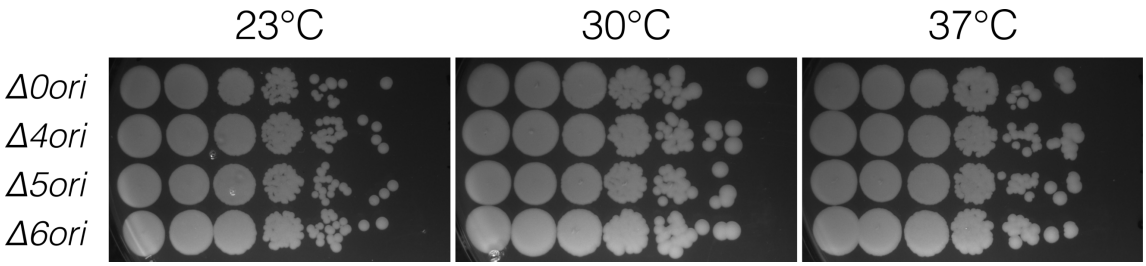


Figure 4.14 Deleting up to six active origins on chromosome VI does not impact on colony growth. *A series dilution assay was carried out with the most concentrated cells at 10^7 cells/ml and serial ten-fold dilutions. Cells were grown on YPAD plates at the temperatures indicated for two days prior to imaging.*

Subsequently, a mean generation time experiment for ChrVI replicon strains was carried out at 30 °C in YPAD medium. Cells from *Δ0ori, Δ4ori, Δ5ori* and *Δ6ori* strains were grown in YPAD medium and readings were taken over four hours. There was no difference in mean generation time between *Δ0ori, Δ4ori, Δ5ori* and *Δ6ori* strains; all strains had a mean generation time of 74 – 78 minutes (Figure 4.15). However, it is known wild-type *S. cerevisiae* have longer mean generation times lower temperatures such as 23 °C, relative to growth at 30 °C (Walsh & Martin 1977). Therefore, mean generation times were determined

using cultures grown at 23 °C as a longer mean generation time allowed for collection of more OD_{600} readings whilst cells remained in mid-log phase. It was found that all strains had a mean generation time of approximately 135 minutes at 23 °C, and no significant differences (using a threshold of $p < 0.05$) were observed between strains with a large replicon and wild-type cells (Figure 4.16).

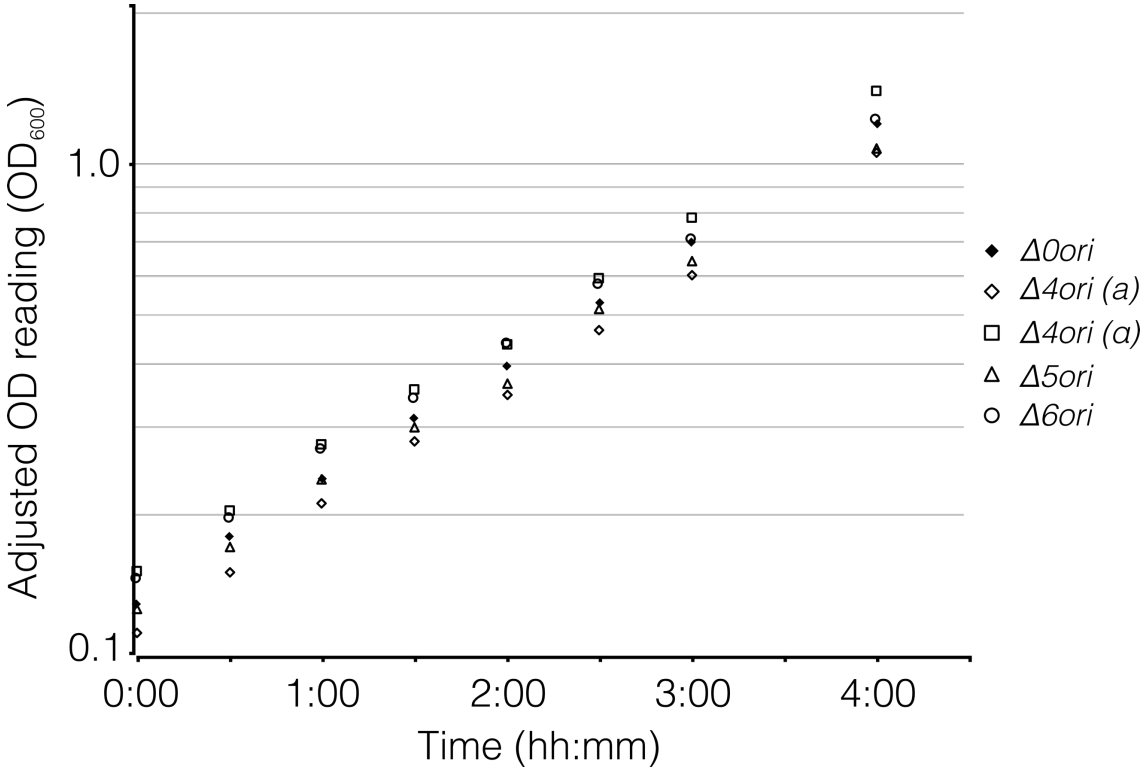


Figure 4.15 Growth curves at 30 °C for strains with origin inactivations on ChrVI. *OD_{600} readings from yeast grown in liquid YPAD at 30 °C were taken over four hours and plotted against time. The gradient of the line of best fit (not plotted) indicates rate of growth.*

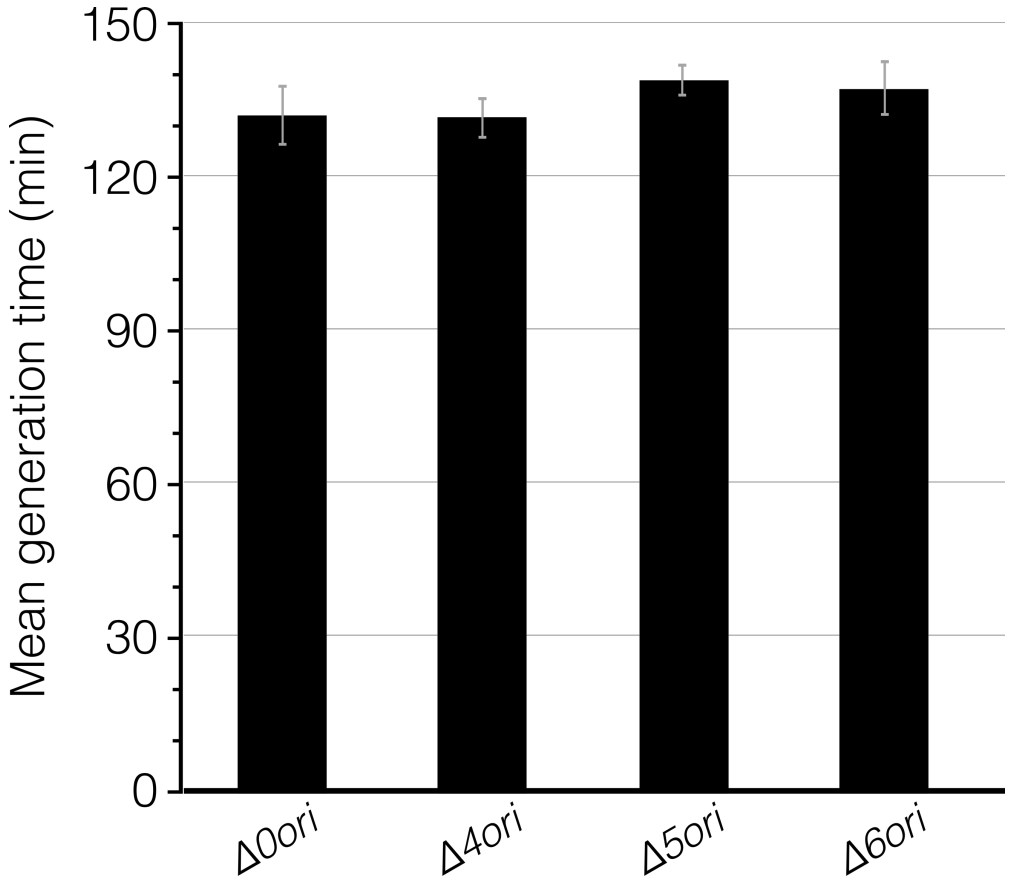


Figure 4.16 Deleting up to six active origins on chromosome VI does not impact on mean replication time. *Mean replication time was calculated from growth curve experiments in liquid YPAD at 23 °C. Four to five repeats per strain were carried out, and error bars represent one standard deviation. No significant differences were found between any large replicon strain and wild-type (Δ0ori) growth (Student's t-test, two-tailed) using a $p < 0.05$ threshold with Bonferroni correction (3 comparisons, adjusted p value = 0.0167).*

YPAD is a rich medium which contains extra nutrients relative to the minimal –HIS medium used for the ChrIII disome strains. This study wanted to preclude the possibility that the medium used resulted in the difference observed between these two systems in the impact a large replicon had on mean

generation time. Therefore, the mean generation time experiment was repeated in complete minimal medium (Figure 4.17) which had the same composition as -HIS medium with the addition of histidine.

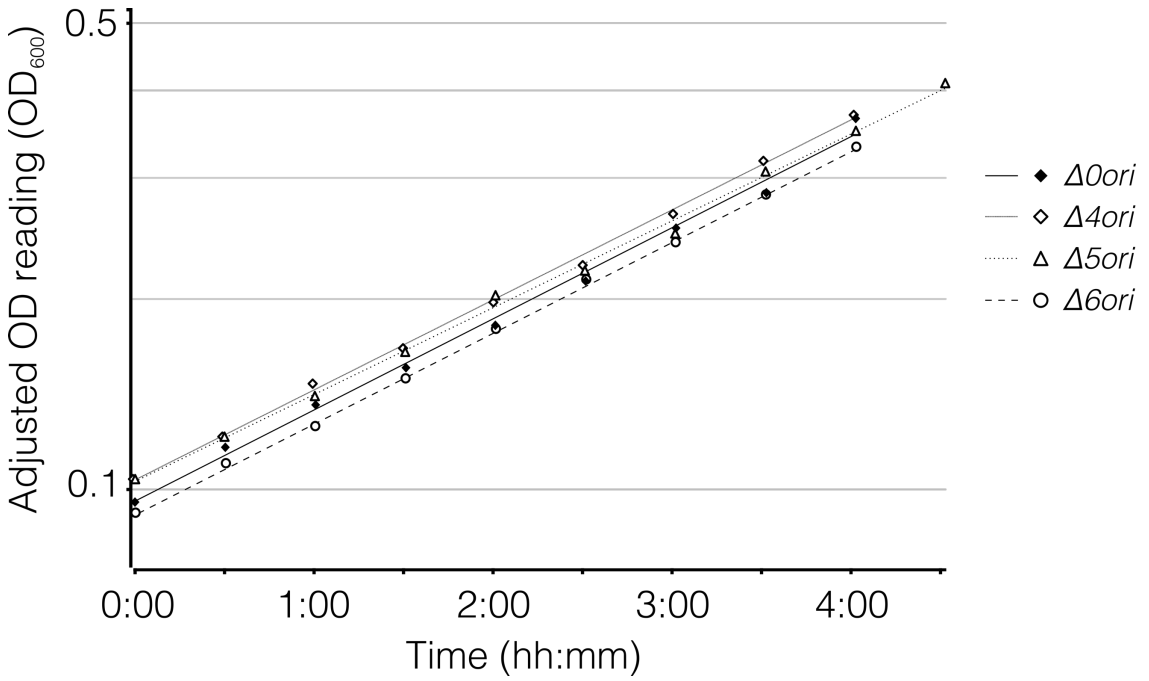


Figure 4.17 Example growth curves for strains with ChrVI origin inactivations in complete medium. *OD_{600} readings from yeast grown in liquid complete minimal medium at 25 °C were taken over five hours and plotted against time. The gradient of the line of best fit (see key) indicates rate of growth.*

Cells from *Δ0ori, Δ4ori, Δ5ori* and *Δ6ori* strains were grown at 25 °C in complete medium and OD_{600} readings were taken over five hours (Figure 4.17). All strains had a mean generation time of approximately 135 – 140 minutes, the same as that observed at a similar temperature in rich YPAD medium (Figure 4.18). This showed that the difference in growth rate between origin deletion and wild-type strains was not due to the medium used.

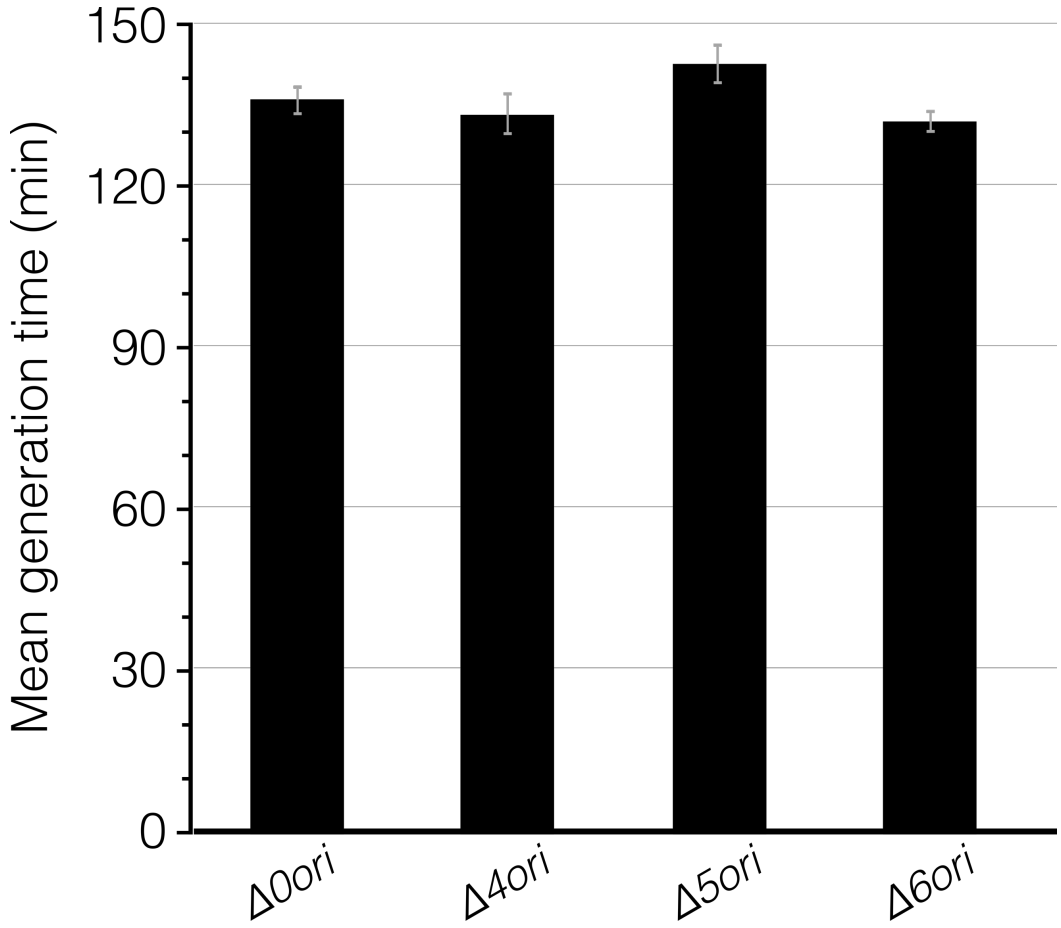


Figure 4.18 The mean growth rates of* Δ4ori, Δ5ori *and* Δ6ori *strains are indistinguishable from that of* Δ0ori *in complete medium. *Mean replication time was calculated from growth curve experiments in liquid complete medium at 25 °C. Three to five repeats per strain were carried out, and error bars represent one standard deviation. No significant differences were found between any large replicon strain and wild-type (*Δ0ori*) growth (Student's t-test, two-tailed) using a p < 0.05 threshold with Bonferroni correction (three comparisons, adjusted p value = 0.0167).*

Conversely, the greatest difference in colony growth between Δ*5ori* and Δ*0ori* ChrIII fragment strains was seen at 37 °C (Figure 4.3). However, since no difference in ChrVI replicon strain growth rate was observed at 23 °C (Figure 4.16), 25 °C (Figure 4.18) or at 30 °C (Figure 4.15), it is unlikely that a difference would be seen at a higher temperature. Therefore, the difference in growth rate between Δ*5ori* and Δ*0ori* ChrIII fragment strains (Figure 4.5) was not due to the medium used or the temperature at which the growth rate was measured.

4.6 DNA replication is delayed at inactivated ACS sites

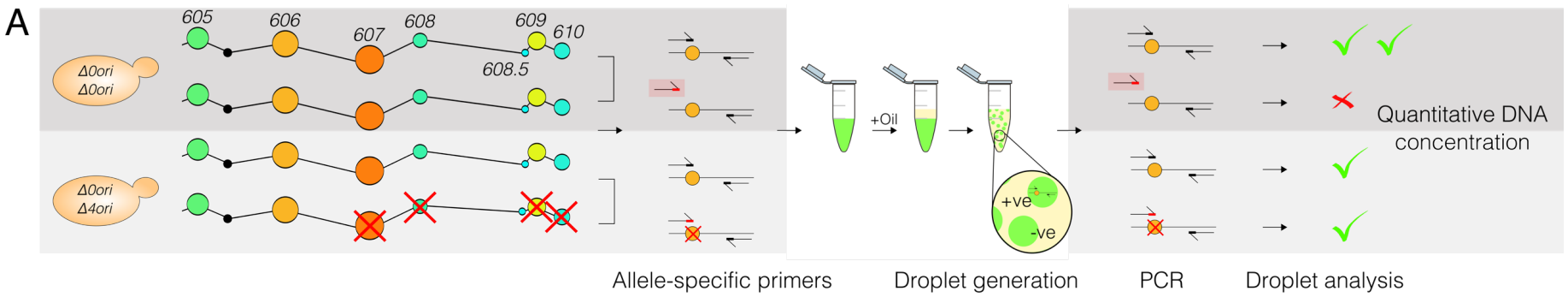
It was shown previously in this chapter (4.2.2) that DNA replication of a chromosome III disome without active origins was delayed to late S phase. This was in contrast to the early replication of chromosome III in wild-type *S. cerevisiae* cells (Figure 4.6). Therefore, this study aimed to examine the replication of the large replicon on the right arm of chromosome VI in Δ*4ori,* Δ*5ori* and Δ*6ori* strains. Initially, diploid yeast cells were used which were heterozygotes for Δ*0ori***,** Δ*4ori*. Using heterozygote yeast cells ensured the wild-type control was at the same cell cycle stage as the origin deletion allele (an internal control). A diploidΔ*0ori* strain was used as a second external control to assess whether the internal control was required.

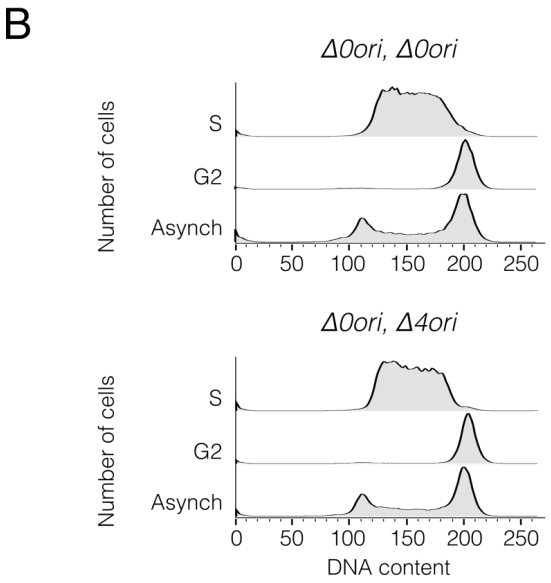
Droplet Digital PCR (ddPCR) was used to measure normalised copy number of discrete loci. This technique was selected because it allowed the use of allele-specific primers which would only amplify from either the Δ*0ori* or the Δ*4ori* ACS site (Figure 4.19A). Additionally, ddPCR has been shown to be capable of distinguishing small differences in replication timing (Batrakou *et al.* 2018). Asynchronous cells were sorted by FACS (Figure 4.19B) to collect S

phase (replicating) and G2 phase (non-replicating) fractions of cells. The flow cytometry profiles of ∆*0ori*, ∆*0ori* and ∆*0ori*, ∆*4ori* cells were indistinguishable, indicating ∆*0ori*, ∆*4ori* cells could pass through the cell cycle with kinetics similar to those of wild-type diploid yeast (Figure 4.19B).

Genomic DNA was extracted from sorted cell fractions and digested using restriction enzyme to smaller fragments. The enzyme selected did not cut within the sites of interest (the bases surrounding *ARS607*, *ARS608*, *ARS609* or *ARS610*). PCR reactions were set up using 0.5 ng digested genomic DNA and 1x EvaGreen Digital PCR SuperMix (Bio-Rad) alongside primers designed to amplify regions approximately 200 bp surrounding each origin site. Samples were processed using a droplet generator (Bio-Rad QX200) and then a PCR reaction was performed. After PCR, DNA concentration was determined using a droplet reader and mean concentrations of DNA in arbitrary units, from 3 technical repeats, were generated, alongside maximum and minimum error values, by Bio-Rad Quantasoft software. Early and late replicating control regions were also amplified as controls, and each region of interest was normalised to controls. This resulted in a normalised range of DNA copy number between 1n (early) and 2n (late) (Figure 4.19C).

Both diploid wild-type cells and the wild-type allele in the ∆*0ori*, ∆*4ori* heterozygote had copy numbers at ACS sites which were in agreement with one another and with published sort-seq data of a wild-type yeast strain (Figure 4.19C) (Müller *et al.* 2013). Hence, it was concluded that the external control of sorting S phase cells from a wild-type yeast strain was suitable for future experiments.


A
605
606
607
608
609
610
608.5
Δ0ori
Δ0ori
Δ0ori
Δ4ori
+Oil
+ve
-ve
Quantitative DNA concentration
Allele-specific primers
Droplet generation
PCR
Droplet analysis


B
Δ0ori, Δ0ori
Number of cells
S
G2
Asynch
0
50
100
150
200
250
Δ0ori, Δ4ori
Number of cells
S
G2
Asynch
0
50
100
150
200
250
DNA content

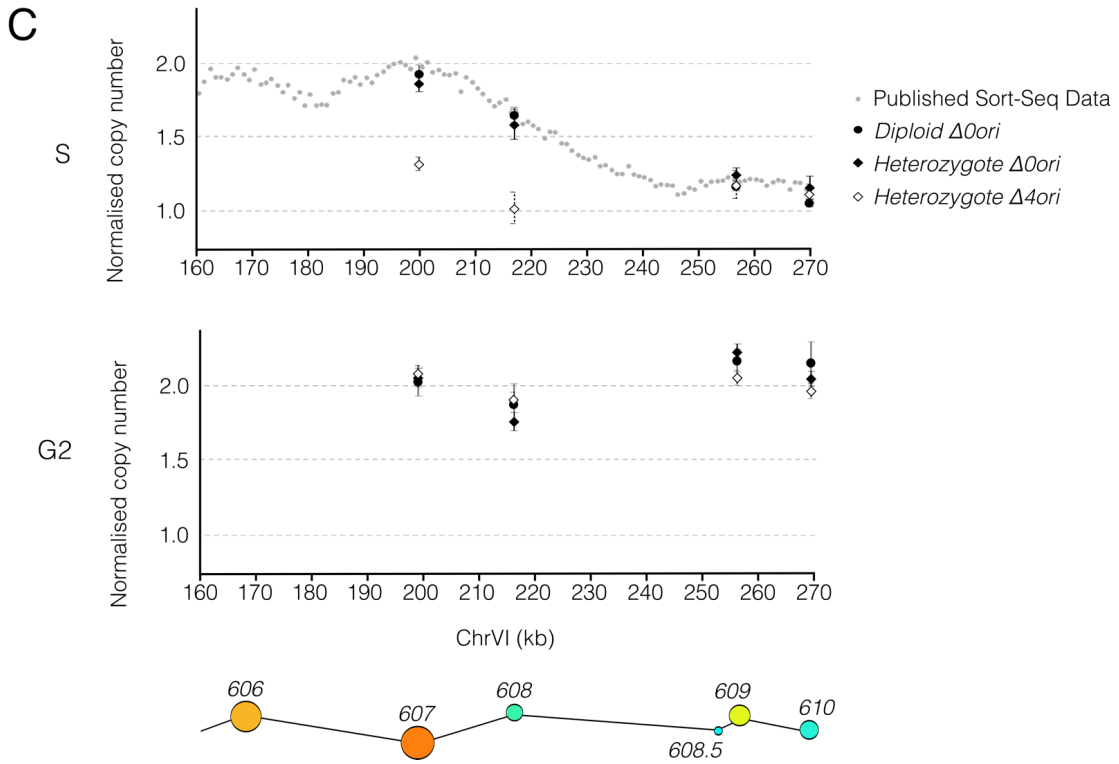

C
S
Normalised copy number
2.0
1.5
1.0
160
170
180
190
200
210
220
230
240
250
260
270
Published Sort-Seq Data
Diploid Δ0ori
Heterozygote Δ0ori
Heterozygote Δ4ori
G2
Normalised copy number
2.0
1.5
1.0
160
170
180
190
200
210
220
230
240
250
260
270
ChrVI (kb)
606
607
608
608.5
609
610

Figure 4.19 Localised DNA replication is delayed when multiple active origins are inactivated. *S or G2 phase cells from diploid yeast were enriched for by FACS, and DNA was extracted. DNA was used in allele-specific droplet-digital PCR to measure relative replication of wild-type and inactivated ACS sites. Chromosome VI is represented as in Figure 4.12. (A) Schematic of allele-specific droplet-digital PCR. Mutant primers (arrow with red box) are unable to bind to wild-type ACS sequence, indicated by a red box surrounding the primer. (B) FACS profiles, showing S and G2 phase cell enrichments, and the asynchronous population for (Top) diploid* ∆0ori *and (Bottom) heterozygote* ∆0ori*,* ∆4ori *cells. (C) Normalised copy number of diploid wild-type and heterozygote ACS sites amplified by PCR. Published sort-seq data is shown as a replication timing profile for comparison (Müller et al. 2013). (Top) S phase DNA (Bottom) G2 phase control DNA.*

The ∆*4ori* allele showed reduced normalised copy number in S phase cells at both *ars607* and *ars608* alleles relative to wild-type *ARS607* and *ARS608* alleles (Figure 4.19C). At both ACS sites the reduction in normalised copy number was greater than 0.5, representing a delay of more than half the length of S phase. However, in the G2 cells the ACS sites for both ∆*0ori*, ∆*0ori* and ∆*0ori*, ∆*4ori* were present in equal normalised copy numbers. This indicated that the delay to replication of the replicon in S phase did not prevent replication from occurring prior to cell division.

No delay in replication was observed during S phase at *ars609* or *ars610* alleles (Figure 4.19C). However, the observed late replication timing of both *ARS609* and *ARS610* alleles indicated that any difference would be hard to detect

in pooled S phase cells. The observed late replication timing of *ARS609* and *ARS610* has been previously reported; *ARS609* has a mean replication time of 32 minutes (Yabuki *et al.* 2002) and *ARS610* has not been identified by any large scale timing study which suggests it is a dormant origin, not typically active during S phase (Alvino *et al.* 2007; Raghuraman *et al.* 2001; Yabuki *et al.* 2002). Additionally, *ARS610* is found within the X element in the telomere, a repetitive DNA element which is hard to study by genomic methods which map sequenced DNA to a reference genome.

A 'G2-early' fraction of cells between the S phase and G2 phase (hence forth G2-late) fractions was previously reported to resolve DNA replication timing of late-replicating regions when used in a technique termed G2-seq (Figure 4.20) (Foss *et al.* 2017). Cells in this fraction have completed bulk DNA replication but some are still completing replication of the latest replicating regions, leading to a larger dynamic range of DNA replication only in late-replicating regions (the majority of the genome exists at a copy number of two in these cells).

Therefore, in order to investigate *ARS609* and *ARS610* S phase, G2-early and G2-late fractions of cells were sorted using FACS. As expected, flow cytometry profiles showed a slight difference in DNA content between G2-early and G2-late cells (Figure 4.21A). Haploid ∆*0ori* and ∆*4ori* strains were used, as the 'external' control had been shown to be comparable to the 'internal' control (Figure 4.19C).

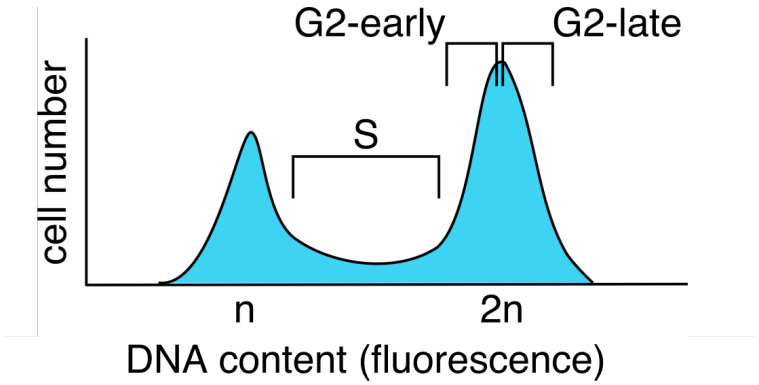


Figure 4.20 In G2-seq an extra fraction of cells is sorted. *The G2-early fraction is sorted from cells which would normally not be included in S phase or G2 (G2-late) phase fractions.*

Replication was delayed in S phase at *ars607* and *ars608* in the ∆*4ori* strain to a greater extent than that observed in the diploid ∆*0ori*, ∆*4ori* strain (compare Figure 4.21B and 4.19C). However, both the G2-early and G2-late fractions of cells showed that DNA replication was completed in both the ∆*0ori* and ∆*4ori* strains, with normalised copy numbers of approximately two (Figure 4.21B). Therefore, it was concluded that ChrVI replication is complete in the G2-early fraction of cells in both the ∆*0ori* and ∆*4ori* strain, making the G2-early fraction of cells unsuitable to examine for the examination of the impact of inactivating the late-replicating *ARS609* and ARS*610* origins.

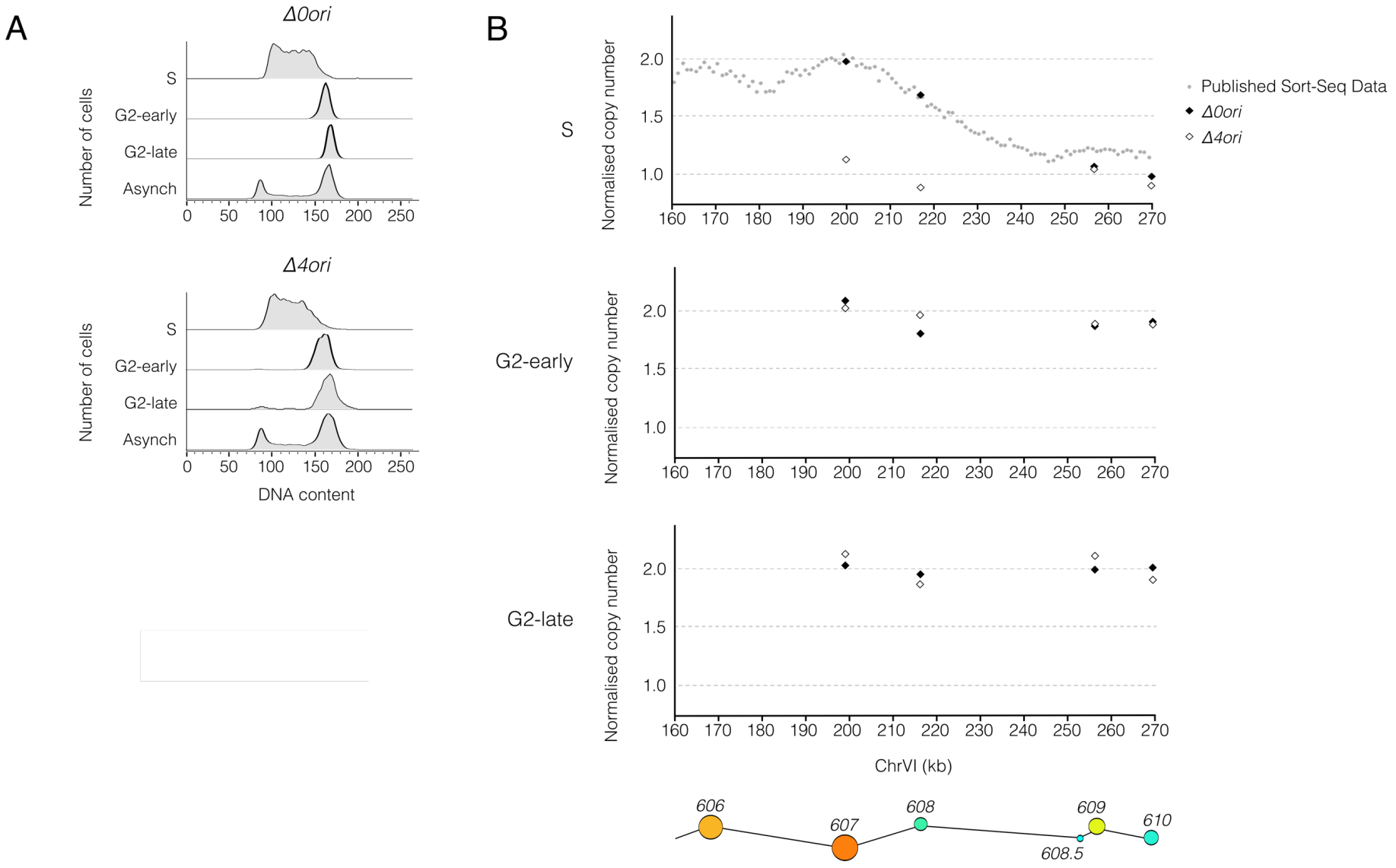

A
Δ0ori
Number of cells
S
G2-early
G2-late
Asynch
0 50 100 150 200 250
Δ4ori
Number of cells
S
G2-early
G2-late
Asynch
0 50 100 150 200 250
DNA content
B
S
G2-early
G2-late
Normalised copy number
2.0
1.5
1.0
160 170 180 190 200 210 220 230 240 250 260 270
Published Sort-Seq Data
Δ0ori
Δ4ori
ChrVI (kb)
606
607
608
608.5
609
610

Figure 4.21 The G2-early fraction of cells does not resolve the replication timing of late-replicating regions. *S or G2 phase cells from haploid yeast were enriched for by FACS, and DNA was extracted. DNA was used in allele-specific droplet-digital PCR to measure relative replication of wild-type and inactivated ACS sites. Chromosome VI is represented as in Figure 4.12. (A) FACS profiles, showing S, G2-early and G2-late phase cell enrichments, and the asynchronous population for (Top)* Δ0ori *and (Bottom)* Δ4ori *cells. (B) Normalised copy number of Δ0ori and Δ4ori ACS sites amplified by PCR. Published sort-seq data are shown as a replication timing profile for comparison (Müller et al. 2013). (Top) S phase DNA (Middle) G2-early phase DNA (Bottom) G2-late phase DNA.*

4.7 High-resolution examination of DNA replication in S phase

An alternative approach to gain higher resolution of DNA replication around *ars609* and *ars610* was to examine the replication timing across the entire replicon. Therefore, this study used sort-seq to further examine the replication of the large replicon in the ChrVI *Δ4ori, Δ5ori* and *Δ6ori* strains. Cells were grown asynchronously at 30 °C and harvested in mid-log phase for cell sorting by FACS. Genomic DNA from sorted S phase and G2 phase cells was subjected to high-throughput sequencing to determine the relative S phase copy number (a proxy for replication time). Haploid strains were used since high-throughput sequencing would not be able to distinguish between the wild-type and mutant chromosome (e.g. in a heterozygous diploid) due to the lack of SNPs between the mutated *ACS* sites.

S phase (replicating) fractions were normalised using non-replicating fractions to reduce noise (Figure 4.22). Typically, the G2-late fraction would be used for normalisation during sort-seq (Müller & Nieduszynski 2012). However, inspection of the raw data across the late-replicating right arm of indicated a reduced copy number (reads per bin) in G2-late samples, even in ∆*0ori* yeast cells (Figure 4.22, ∆*0ori,* ∆*4ori,* ∆*5ori*). This was consistent with an observed late replication time for telomeres across the genome. Therefore, three alternative non-replicating samples were considered: genomic DNA from G1 phase cells, alpha factor-arrested cells and stationary phase cells. Genomic DNA from both alpha factor-arrested cells and stationary phase cells showed uniform copy number across the right arm of ChrVI in the ∆*5ori* strain (Figure 4.22, ∆*5ori*).

Stationary phase cells were selected as an appropriate non-replicating control for sort-seq of origin deletion strains. The ∆*0ori,* ∆*4ori* and ∆*5ori* replicating S phase samples were normalised to the ∆*5ori* stationary phase sample. All of these samples used the same number of PCR amplification steps during the preparation of Next-seq DNA libraries. The ∆*6ori* S phase sample required a different number of PCR amplification steps during Next-seq library preparation. This was due to a lower yield of DNA after DNA extraction and sonication, meaning less DNA was used as input in for making Next-seq libraries. Therefore, the ∆*6ori* S phase sample was normalised to a separate, ∆*6ori* stationary phase sample which was prepared using the same number of PCR amplification steps (Figure 4.22, ∆*6ori*). The normalised sort-seq S phase samples are plotted in Figure 4.23.

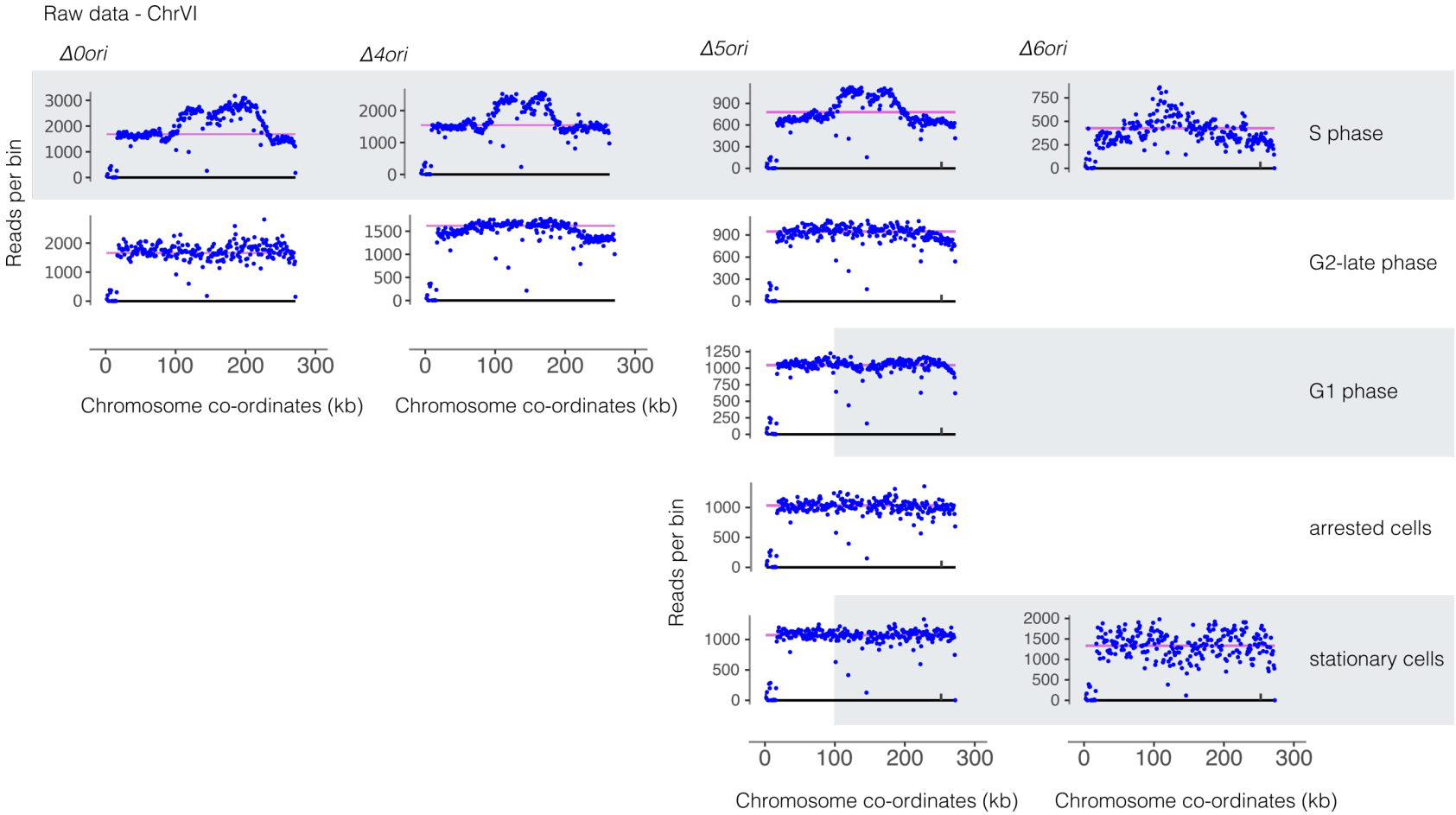

Raw data - ChrVI
Δ0ori
Δ4ori
Δ5ori
Δ6ori
Reads per bin
S phase
G2-late phase
G1 phase
arrested cells
stationary cells
Chromosome co-ordinates (kb)
Chromosome co-ordinates (kb)
Chromosome co-ordinates (kb)
Chromosome co-ordinates (kb)

Figure 4.22 Stationary phase cells are appropriate for normalisation of Sort-seq data. *Sort-seq data are shown as reads per bin across ChrVI, which is not normalised (raw data). Genomic DNA from S phase (*∆0ori*,* ∆4ori*,* ∆5ori *and* ∆6ori*), G2 phase (*∆0ori*,* ∆4ori *and* ∆5ori*), G1 phase (*∆5ori*), alpha factor-arrested cells (*∆5ori*), and stationary phase cells S phase (*∆5ori *and* ∆6ori*) was subjected to high-throughput sequencing and mapped to the* S. cerevisiae *reference genome to determine the number of reads mapping per bin.*

Replication of the large replicon was delayed in the ∆*4ori* strain (Figure 4.23). Relative copy number of both *ars607* and *ars608* was reduced relative to the wild-type ∆*0ori* relative copy number. As in ddPCR (Figure 4.19C), the resolution of the late-replicating DNA surrounding *ars609* and *ars610* meant it was not possible to distinguish a difference between the replication of this region in the ∆*0ori* and ∆*4ori* strains. However, there were no clear peaks indicating replication initiation throughout the large replicon. The *ARS606* peak observed in the ∆*4ori* strain was at the same height as that seen in the ∆*0ori* strain. Relative copy number to the right of this peak decreased to approximately 1.0 at *ars607* and remained low throughout the large replicon. This suggested replication of the large replicon was dependent on replication forks moving rightwards from *ARS606* in the ∆*4ori* strain (Figure 4.23).

Relative copy number of the ∆*5ori* strain appeared to be very similar to that in the ∆*4ori* strain (Figure 4.23). The relative copy number across the large replicon appeared to be flatter than in the ∆*4ori* strain, which might have been due to the small effect of active *ARS608.5* in the ∆*4ori* strain. However, the raw S phase data for both ∆*4ori* and ∆*5ori* strains showed no peak in the reads per

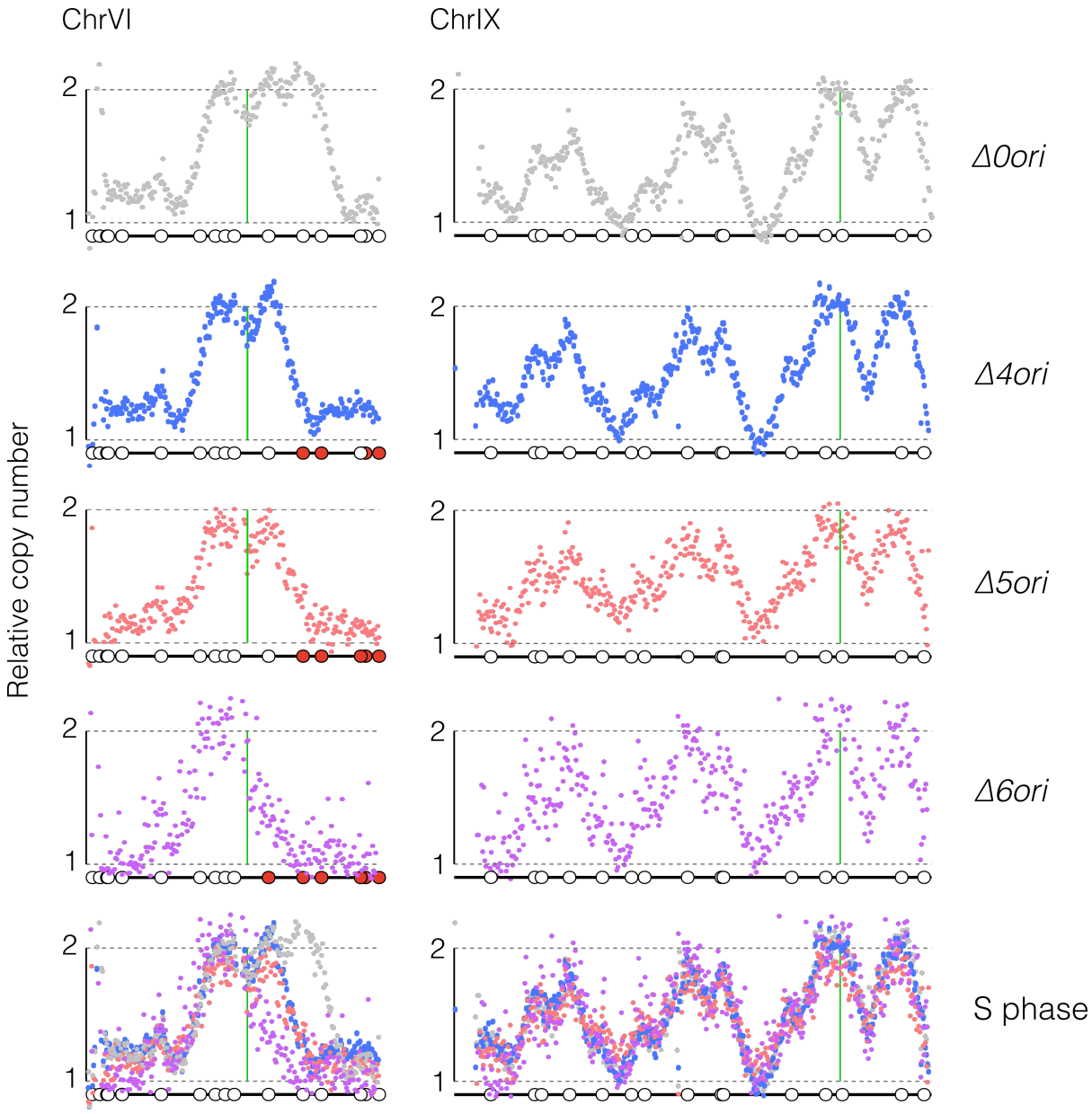


Figure 4.23 DNA replication across a large replicon on chromosome VI is delayed when multiple active origins are inactivated. *S or G2 phase cells from haploid yeast were enriched for by FACS, and DNA was extracted. Sort-seq was performed to measure relative copy number of the genome in haploid* Δ0ori, Δ4ori, Δ5ori *and* Δ6ori *strains. The length of the chromosome is shown on the X axis as a black line with origins (white circles) and inactive origins (red circles). The centromere is shown by a vertical green line. ChrIX is shown as a control.*

bin across the large replicon (Figure 4.22, Δ*4ori,* Δ*5ori*), indicating that the slight increase observed in the normalised sort-seq Δ*4ori* data could be due to a slightly reduced copy number across this region in the non-replicating control sample, to which the data were normalised.

No peak in relative copy number was observed at the *ARS606* locus of the Δ*6ori* strain, indicating the inactivation of *ARS606* to *ars606* (Figure 4.23). In the presence of this 135-kb replicon, no peaks of initiation across the replicon were observed, which suggested that replication was now dependent on replication forks which initiated at the neighbouring *ARS605* or origins to the left of *ARS605*.

4.8 Summary

Having previously examined the regulation of DNA replication origins (Chapter 3), this study aimed to examine the consequences of inactivating DNA replication origins. Particularly, this chapter examined the impact of a large replicon on the growth and DNA replication of *S. cerevisiae* cells. It has been shown previously that *S. cerevisiae* cells are able to maintain an origin-less fragment of chromosome III (Dershowitz *et al.* 2007). Here, the mean generation time of a yeast strain with a second ChrIII 174 kb fragment was found to be increased when five active origins on the fragment were inactivated (Figure 4.4, 4.5).

This finding was supported by the first origin-less fragment maintenance (*ofm*) mutant: *rad9* (Theis *et al.* 2007). *RAD9* is a DNA damage checkpoint component, required to arrest cell cycle division in G2 after DNA damage (Weinert & Hartwell 1988). It has been shown that *RAD9* is essential for maintaining an origin-less fragment of ChrIII and that, in its absence, this fragment

was lost at a highly increased rate (Theis *et al.* 2007). Rad9 has recently been shown to play a role in a late cell cycle checkpoint in yeast; it is required to prevent cytokinesis prior to DNA replication completion (Ivanova *et al.* 2018). In this study the 174 kb ChrIII fragment included dormant origins *ARS300* – *ARS304* (Figure 4.1). However, these origins are late-firing and DNA replication was found to be delayed until after the majority of the genome had fully replicated (Figure 4.6). The presence of *RAD9* would facilitate late DNA replication progression from dormant origins by arresting mitosis until DNA replication completed. Therefore, this study suggested that in yeast with a second copy of a 174 kb ChrIII fragment, DNA replication is delayed when active origins are removed from the fragment. These data were consistent with the delay to DNA replication leading to activation of the Rad9 dependent cell cycle checkpoint during anaphase. The Rad9 checkpoint is activated by CDK-dependent Rad9 phosphorylation (Wang *et al.* 2012) and functions during anaphase to delay nuclear division during mitotic DNA synthesis (Ivanova *et al.* 2018). Rad9 is capable of binding Dpb11, which may suggest ongoing replication forks are able to directly or indirectly recruit and activate Rad9, although Dpb11 binding is not required for Rad9 activation in anaphase phosphorylation (Wang *et al.* 2012). Activation of the Rad9 dependent checkpoint would result in a longer cell cycle, consistent with this study's observation of a greater mean generation time.

Subsequently, this study further examined the consequences of a large replicon on cell growth and DNA replication using a genomic model. Three yeast strains were generated (Δ*4ori*, Δ*5ori* and Δ*6ori*) with either 100 kb or 135 kb replicons on the right arm of chromosome VI. During creation of these strains, a potential ACS for the confirmed origin *ACS608.5* was identified (Figure 4.9, 4.10).

Mutation of this ACS site led to reduced origin activity (Figure 4.12), indicating that the potential ACS site identified is the primary ACS site for *ACS608.5*.

This study found no effect on mean generation time when 100 kb or 135 kb replicons were generated on the right arm of chromosome VI (Figure 4.15, 4.16). This result was consistent with a lack of observed growth defect when deleting seven active origins on the left arm of chromosome VI (Bogenschutz *et al.* 2014) and previous work showing that following the inactivation of *ARS606*, *ARS607* and *ARS608*, growth rates are similar to wild-type yeast (Müller 2012). Additionally, it has been shown that triple *ars606*, *ars607* and *ars608* mutant yeast are not more sensitive to either the DNA damaging agent mechlorethamine or UV radiation than wild-type cells (Müller 2012). This suggested that the presence of a large replicon itself does not impair cell growth.

The difference between the absence of a growth defect observed in the genomic ChrVI models and the growth defect observed in the strain with a second ChrIII fragment may be explained by the presence of early firing origins present on ChrVI. In the ∆*4ori* and ∆*5ori* strains, *ARS606* was still active, whilst in the ∆*6ori* strain *ARS605* remained functional and in the published ∆*7ori* strain *ARS607* was active (Bogenschutz *et al.* 2014). The presence of early firing origins may have allowed completion of DNA replication within the normal duration of S phase. Evidence in support of this is seen in the ddPCR data, where in G2-early and G2-late fractions, the normalised copy number of the ACS sites could not be distinguished between ∆*0ori* and ∆*4ori* strains (Figure 4.21). This is despite a clear delay in replication of the ACS sites of ∆*4ori* cells being observed during S phase (Figure 4.19, 4.21). In comparison, at the 90-minute time point equivalent for the ChrIII fragment strain, despite the rest of the genome being fully replicated,

(representing a G2 population of cells) the majority of the ChrIII fragment is at a relative copy number of 1.4 (Figure 4.6). Therefore, the different impact a large replicon had on mean generation time between a strain with a second ChrIII fragment and a genomic ChrVI large replicon could result from a difference in the extent of the delay caused to DNA replication. It is proposed that one reason the ChrVI replicon replicates earlier is due to active, early-firing neighbouring origins.

One caveat to the replication delay seen in the strain with a second ChrIII fragment is that the ChrIII fragment is also at a relative copy number of 1.4 at the 50-minute time point equivalent (Figure 4.6). This is prior to a peak in relative copy number near the dormant origins, suggesting they have not activated at this time point. The increase in copy number seen across the ChrIII fragment suggests some unknown initiation sites are activating on ChrIII. This low-level replication in the absence of clear origin activation was not observed in the ChrVI replicon strains. This further difference between the ChrIII model and the ChrVI model used to investigate large replicons could be due to the early-activating origins on ChrVI masking any low-level initiation from unknown sites, or due to the synchronous cell cycle release being able to resolve low-level replication which is not observed in sort-seq due to the average of cells in all stages of S phase being used.

Although the mean generation time of ∆*4ori*, ∆*5ori* and ∆*6ori* cells was not increased, DNA replication was delayed. Raw sort-seq data showed that even after bulk DNA replication there was a slight reduction in reads per bin mapping to the large replicon on ChrVI (Figure 4.23). At the far right of ChrVI this effect could also be seen in ∆*0ori* cells, indicating that the far right of ChrVI is a late-replicating region in wild-type cells (Figure 4.23). This is supported by the late

activation of *ARS609* and *ARS610* observed by ddPCR (Figure 4.19, 4.21) and sort-seq (Figure 4.22). However, the reads per bin mapped evenly across the whole of ChrVI in cells G1-phase arrested cells or in stationary (G0) phase cells, indicating that prior to cytokinesis cells do complete replication of the replicon (Figure 4.23). Therefore, it can be speculated that some cells completed replication of the large ChrVI replicon during late G2 or early M phase.

Replication during late G2 has been reported alongside a Rad9 checkpoint which prevents cell cycle progression until this process is completed (Ivanova *et al.* 2018). Although the mean generation time was not extended in the ChrVI replicon strains (suggesting that the Rad9 checkpoint was not activated) this does not preclude some level of DNA replication during late G2 phase occurring, but completing prior to checkpoint activation. Additionally, in some circumstances a delay in one cell cycle stage can be compensated for by the shortening of another cell cycle stage, for example G1 (Muller 1991). Alternatively, as previously noted, the *Δ0ori* cells also had a reduced reads per bin in G2 at the far right of the replicon (Figure 4.23). It could therefore be conjectured that a level of Rad9 activation is occurring in wild-type cells to facilitate replication of ChrVI.

In conclusion, this chapter examined the impact of a large replicon in two yeast models. Both models showed a delay in DNA replication when active origins were removed. In the model with a second copy of a 174 kb ChrIII fragment this correlated with an increased mean generation time, however, this was not observed in the ChrVI model. It was proposed that while a cell cycle checkpoint may be activated in the ChrIII model, the neighbouring active origins on ChrVI could be conjectured to be able to rescue the delayed DNA replication. The far right of ChrVI was shown to replicate late, and it is suggested that this region is

not fully replicated in wild-type or large replicon-containing cells prior to G2. The use of a G2-early fraction of cells did not permit resolution of this late-replicating region, therefore sort-seq was used to examine the replication across the entire replicon. No peaks in replication were observed, indicating that no replication initiation was occurring in early-to-mid S phase. Other ways to visualise and resolve this late-replicating region of chromosome VI were investigated in the following chapter.

5 Examining replication timing and polymerase usage across a large replicon in *Saccharomyces cerevisiae*

5.1 Introduction

In Chapter 4 it was shown by sort-seq and ddPCR that the right arm of chromosome VI, in the region surrounding *ARS608.5*, *ARS609* and *ARS610*, was at a copy number of one in S phase cells (Figure 4.19, 4.21, 4.22). This indicated that DNA across large replicons in *S. cerevisiae* was either replicated late during S phase or during the G2 or M phases of the cell cycle. For the discussion of these results, G2 phase is defined as the phase where the majority of DNA replication has been completed and M phase is defined as the process of chromosome segregation having already begun. It has been shown in both human and yeast cells that DNA replication can occur during the early stages of mitosis (Bhowmick *et al.* 2016; Ivanova *et al.* 2018).

This form of replication is known as MiDAS in human cells (Bhowmick *et al.* 2016). In yeast, the Rad9 checkpoint prevents cells from undergoing mitosis until replication has completed (Ivanova *et al.* 2018; Schwartz *et al.* 2002). In ChrIII disome strains it has been shown that *RAD9* is required for the maintenance of origin-less fragments (Theis *et al.* 2007). Additionally, it was shown by examining the raw sort-seq data that replication of the ChrVI replicon in ∆*4ori* and ∆*5ori* yeast was not completed in the G2 fraction of cells (Figure 4.23). However, the previous chapter also showed that in ChrVI origin mutant strains the cell cycle is not extended (Figure 4.16), suggesting DNA replication may be completed prior to activation of the Rad9 checkpoint. In order to investigate at what time DNA replication occurs across the ChrVI large replicon, synchronous culture

experiments were carried out to enable sequencing of chromosomal DNA from late-S phase cells.

Replication timing data can be supplemented by examining other aspects of replication, such as directionality. In the ChrIII fragment Δ*5ori* strain it was previously found by analysing 2D gels that DNA replication forks across the 174 kb fragment did not exclusively travel away from the dormant replication origins *ARS300-ARS303* (Figure 5.1) (Dershowitz *et al.* 2007). Even at the rightmost probe close to *ARS310*, it was found that 32 % of replication forks were travelling left, towards the dormant origins (Figure 5.1) (Dershowitz *et al.* 2007). Therefore, in order to understand how replication advances across large replicons in *S. cerevisiae* this chapter also examined how the replication fork progressed.

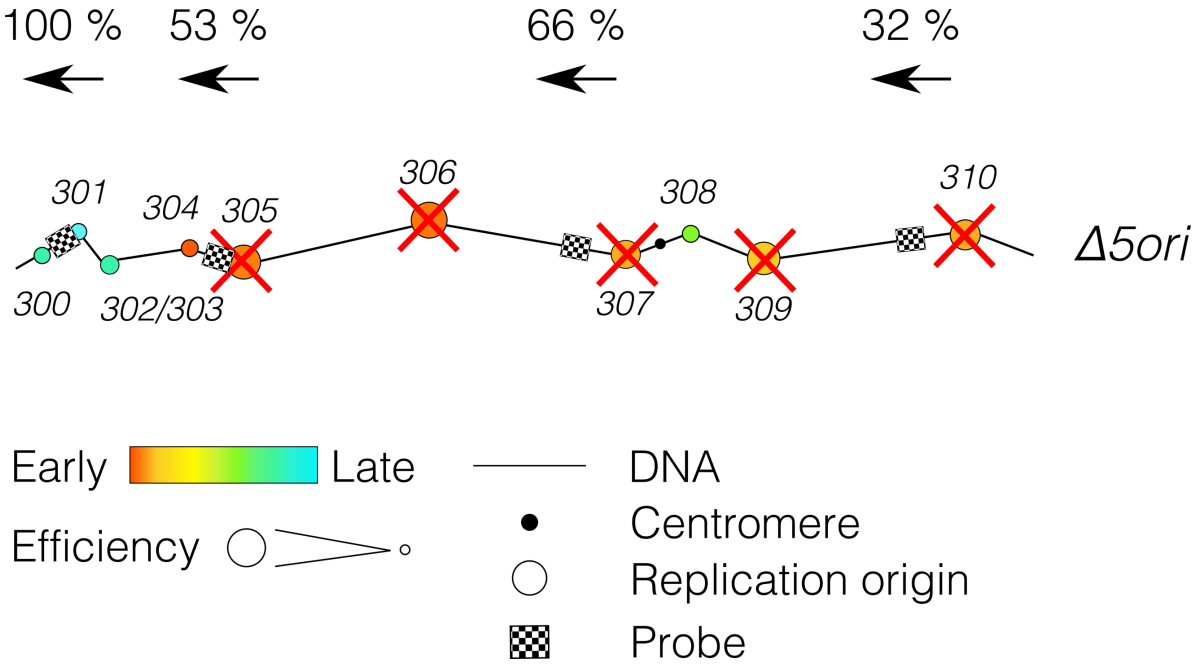


Figure 5.1 Replication forks move in both directions across a large replicon. *Replication origins are represented by circles and labelled with their ARS number (e.g. 305 =* ARS305*). Circle diameter represents efficiency (the mean of origin efficiencies calculated from replication timing data (Hawkins et al. 2013b) and Okazaki fragment data (Smith & Whitehouse 2012) and the colour indicates mean replication time from early (red) to late (blue) based on replication timing data (Hawkins et al. 2013b). A 174 kb fragment of ChrIII with* Δ5ori *is shown, The percentage of replication forks moving leftward is shown by the black arrows, percentages are based on 2D gel analysis (Dershowitz et al. 2007). Figure modified from (Dershowitz et al. 2007).*

A canonical replication fork includes two replicative polymerases; Polε (Pol2) is primarily associated with leading strand synthesis while Polδ (Pol3) is primarily associated with lagging strand synthesis (Johansson & Dixon 2013; Makarova & Koonin 2013). The catalytic subunits of Polε and Polδ derive from

the *POL2* and *POL3* genes respectively. In 2015 multiple techniques to examine the relative use of these polymerases in a strand-specific manner were published: emRiboSeq (Reijns *et al.* 2015), HydEn-seq (Clausen *et al.* 2015; Williams *et al.* 2015), ribose-seq (Koh *et al.* 2015; Storici *et al.* 2015) and Pu-seq (Keszthelyi *et al.* 2015). These methods rely on the use of polymerase mutants with a higher propensity than wild-type polymerases to incorporate ribonucleotides into the DNA (Table 5.1). Modification of key amino acids in the catalytic site of Polε and Polδ change their specificity for deoxyribonucleic acids, leading to either decreased or increased specificity (Clausen *et al.* 2015; McElhinny *et al.* 2010b) (Table 5.1).

Table 5.1 ribonucleotide incorporation rates of mutant polymerases. *Rate of ribonucleotide incorporation into DNA is shown. *Rate inferred from alkali-sensitivity relative to Pol2/Pol3.*

POLYMERASE	*dNTP:rNTP BASE RATIO*	*REFERENCE*
Pol2 (Polε)	2000	(Keszthelyi *et al.* 2015)
Pol2-M644G	91	(McElhinny *et al.* 2010b)
Pol2-M644L	6000*	(McElhinny *et al.*, 2010)
Pol3 (Polδ)	5000	(Keszthelyi *et al.* 2015)
Pol3-L612G	500*	(Lujan *et al.* 2013)

Incorporation of ribonucleotides into DNA would normally be repaired by Ribonucleotide Excision Repair (RER) (Wallace & Williams 2014). Therefore, the RNaseH1 gene (*RNH201*), which is essential for RER, was also deleted (Wallace & Williams 2014), resulting in *pol2-M644G* ∆*rnh201* and *pol3-L612G* ∆*rnh201*

yeast strains. These strains selectively incorporate high levels of ribonucleotides on DNA synthesised by one polymerase. For example, a *pol2-M644G Δrnh201* strain incorporates ribonucleotides in areas synthesised by Polε (Figure 5.2).

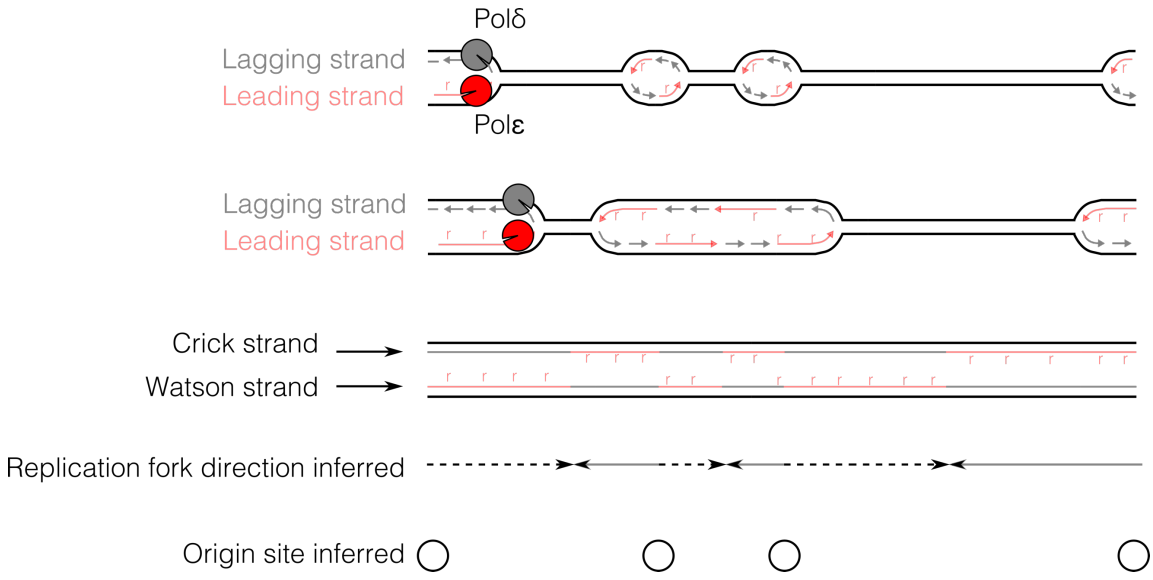


Figure 5.2 Ribonucleotides are incorporated into nascent DNA by mutant replicative polymerases. *A schematic representing a mutant leading strand polymerase (Polε) incorporating ribonucleotides. Strand specific incorporation of ribonucleotides can be detected by techniques such as Pu-seq in the resulting DNA (Keszthelyi et al. 2015). Replication fork direction can be inferred by assuming leading strand synthesis indicates a replication fork traversing 5' to 3' on that DNA strand, and points where leading strand synthesis changes from the Crick to the Watson strand can be inferred to be origins of replication. Origin sites are also inferred at the distal ends of the DNA fragment.*

Pu-seq raw data measures the number of ribonucleotides incorporated into a DNA strand (e.g. the Watson strand) by a mutant Polε (*pol2-M644G*). This is

normalised to the sum of ribonucleotides incorporated into the same strand by both mutant Polε (*pol2-M644G*) and Polδ (*pol3-L612G*) strains. A change in signal indicates a change in polymerase usage on the DNA strand. Assuming that Polε predominantly synthesises leading strand DNA, ribonucleotide signal can be used as a proxy for replication forks traversing DNA in the 5' to 3' direction on that strand (Figure 5.2). Correspondingly, sites where ribonucleotide incorporation by Polε increases rapidly on the Watson strand indicate an origin (Figure 5.2). Conversely, sites where ribonucleotide incorporation by Polε reduce on the Watson strand indicate a termination zone, with forks terminating stochastically. The gradient of the change in ribonucleotide incorporation signal represents the breadth the zone over which termination occurs; where two efficient origins activate at similar times in S phase forks may terminate approximately the same place in most cells, leading to a sharp change in signal.

In order to examine the direction of replication forks replicating the right arm of ChrVI in origin deletion strains, this chapter examined polymerase-usage of Polε and Polδ across the large replicon. Pu-seq has been previously published in the fission yeast strain *Schizosaccharomyces pombe*, but required optimisation for use in *S. cerevisiae*. Advice on this process and unpublished *S. cerevisiae* ∆*0ori* Pu-seq data were provided by a collaborator in the Carr laboratory, Katie Ptasińska (Figure 5.3A). An example window of Pu-seq data for ChrVI right is shown in Figure 5.3B. It was predicted that in the ∆*5ori* strain, all replication forks would travel right from *ARS606* towards the telomere (Figure 5.3B) since no initiation sites were observed across the large ChrVI replicon by sort-seq (Figure 4.23).

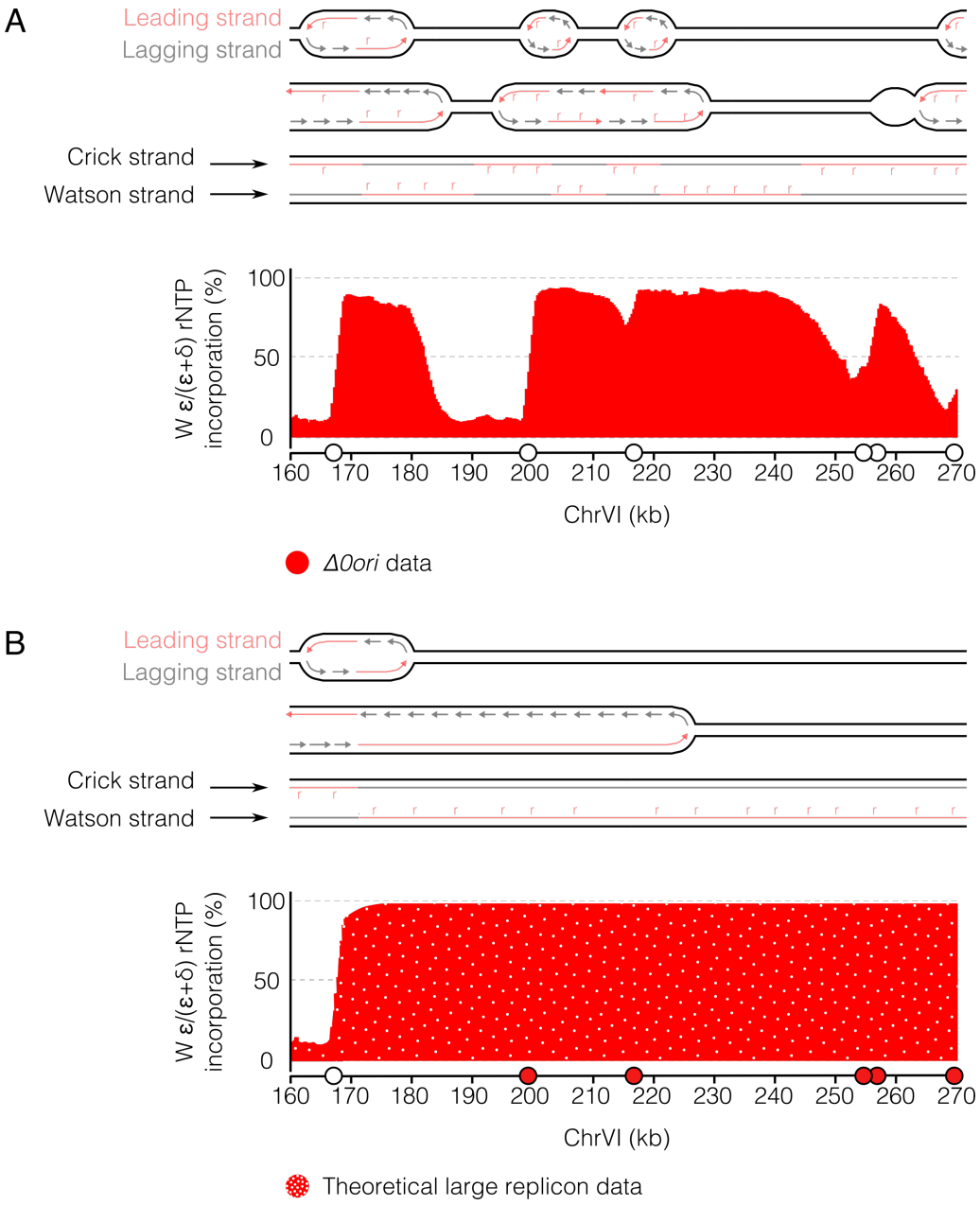


Figure 5.3 A representation of Pu-seq data where all replication initiates at known origins. *In a* pol2-M644G Δrnh201 *strain, ribonucleotides are incorporated into the leading strand. Data is normalised using the sum of ribonucleotides incorporated on the Watson strand from mutant Polε and Polδ strains. (A)* Δ0ori *Pu-seq data from Katie Ptasińska (personal communication). (B)* Δ5ori *Pu-seq theoretical data where it is assumed all replication initiates at* ARS606.

5.2 Replication timing of late-replicating DNA can be resolved using synchronous culture experiments

In order to study the late-replicating right arm of ChrVI synchronous culture experiments were performed. Prior to these experiments, Shin-ichiro Hiraga, a collaborator in the Donaldson laboratory, had created an independent Δ*4ori* strain (HE61) for a separate study (Kedziora *et al.* 2018). HE61 replaced origins with selective markers and is hereafter referred to as Δ*4ori (m)*. A synchronous experiment was performed by Shin-ichiro Hiraga using the Δ*4ori (m)* strain. Cells were arrested in log phase using alpha factor until arrested. Cells were subsequently released from arrest by addition of pronase, passing synchronously through S phase. Samples were taken at time points equivalent to 70 % and 95 % bulk DNA replication, genomic DNA was extracted and sequenced by high-throughput sequencing. Replicating samples were normalised using a genomic DNA sample from non-replicating alpha factor-arrested cells, which were confirmed to have uniform reads per bin across the genome (data not shown).

This experiment showed that when bulk DNA replication was approximately 70 % complete (late S phase) an increased relative copy number was visible at the far right of ChrVI (Figure 5.4, unpublished data). This suggested replication was able to initiate close to the telomere of ChrVI; this will henceforth be referred to as telomeric initiation. Peaks in relative copy number at *ARS606* and other active origins were also observed, indicating replication was occurring canonically on the left arm of ChrVI. At 95 % completion of bulk DNA replication

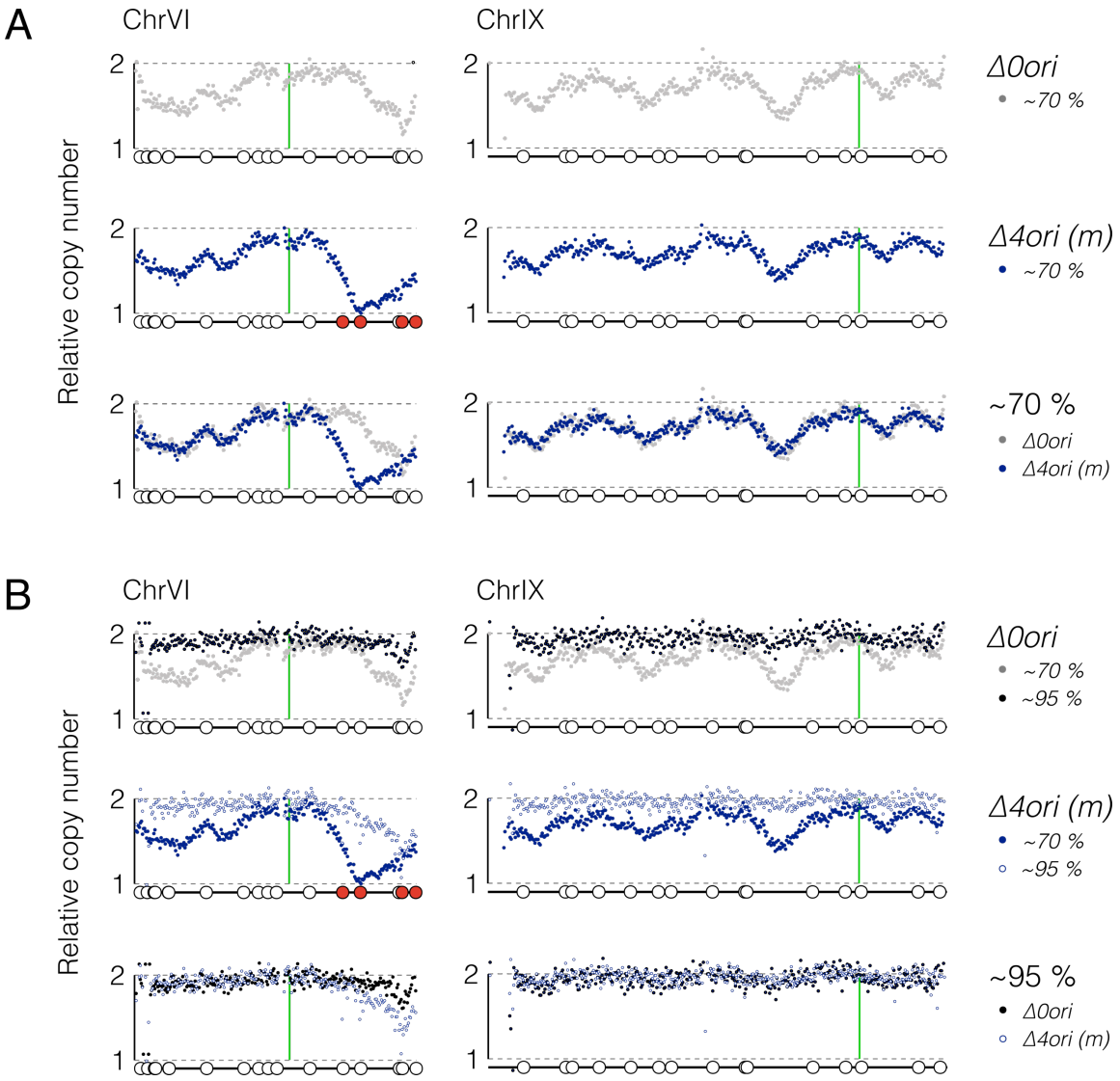


Figure 5.4 Replication of a large replicon on ChrVI right occurs late during S phase and initiates near the telomere. *A synchronous culture experiment was carried out by Shin-ichiro Hiraga (Donaldson laboratory). Relative copy number data for ChrVI and control ChrIX are shown at (A) 70 % and (B) 95 % bulk genome replicated. The 95 % time point data have been published previously (Kedziora et al. 2018). The* Δ4ori (m) *strain HE61 has origins replaced with markers whereas the* Δ4ori *strain ACY162 has origins inactivated by site-directed mutation. Data were normalised to alpha factor-arrested samples.*

the relative copy number of DNA to the right of *ARS606* was close to two in the Δ*4ori (m)* strain (Figure 5.4B). However, the relative copy number at the far right of ChrVI, remained at 50 % This suggested that there was a maximum initiation of 50 % for replication forks initiating in telomeric initiation.

The telomeric initiation observed in the Δ*4ori (m)* strain was not observed in the Δ*4ori*, Δ*5ori* or Δ*6ori* strains by ddPCR or sort-seq (Figures 4.19, 4.21, 4.22). Two explanations were proposed for this discrepancy. First, it was possible that strain differences in the Δ*4ori (m)* strain resulted in this telomeric initiation. Secondly, since sort-seq and ddPCR use the entire fraction of S phase cells which are, on average, 50 % replicated, these techniques may be unable to resolve late S phase initiation. To distinguish between these possibilities, sort-seq was performed on the Δ*4ori (m)* strain (Figure 5.5).

It should also be noted that the normalisation samples used in sort-seq and synchronous culture experiments differed. Stationary phase cells were used to normalise the replicating S phase sample in sort-seq, whereas alpha-factor arrested cells were used to normalise the replicating S phase sample in *the* Δ*4ori (m)* synchronous culture experiment. However, as shown in Figure 4.23, alpha factor-arrested cells were also sequenced during sort-seq for the Δ*5ori* strain. Normalising the Δ*0ori*, Δ*4ori* or Δ*5ori* replicating S phase samples to the alpha factor-arrested cells (data not shown) produced a similar replication timing profile to that produced when the samples were normalised to stationary phase cells (Figure 4.22). The stationary phase cells were selected for normalisation because they had slightly less intrinsic noisiness, producing a clearer plot for visualisation.

Δ*4ori (m)* cells were grown to mid-log phase at 30 °C and sorted by FACS to enrich for S phase replicating cells. Additionally, a non-replicating sample of

stationary Δ*4ori (m)* cells was collected as a control. S phase and stationary phase cells had their genomic DNA extracted and were subject to high-throughput sequencing. Sequencing reads were mapped to the *S. cerevisiae* reference genome. It was confirmed that the stationary phase control had uniform raw reads per bin mapped across the genome (data not shown). The S phase sample was normalised using the stationary phase sample.

When examined by sort-seq the Δ*4ori (m)* strain did not show an increase in relative copy number at the far right of ChrVI (Figure 5.5). Further, the overall pattern of relative copy number seen was similar to that seen in the Δ*4ori* strain. A caveat to this data was the unusual noisiness of the Δ*4ori (m)* samples. However, since the pattern was still distinguishable from that observed in the Δ*4ori (m)* synchronous culture experiment (Figure 5.4) it was concluded that the detection of telomeric initiation was attributable to a difference in the resolution of late-replicating DNA between sort-seq and synchronous culture experiments.

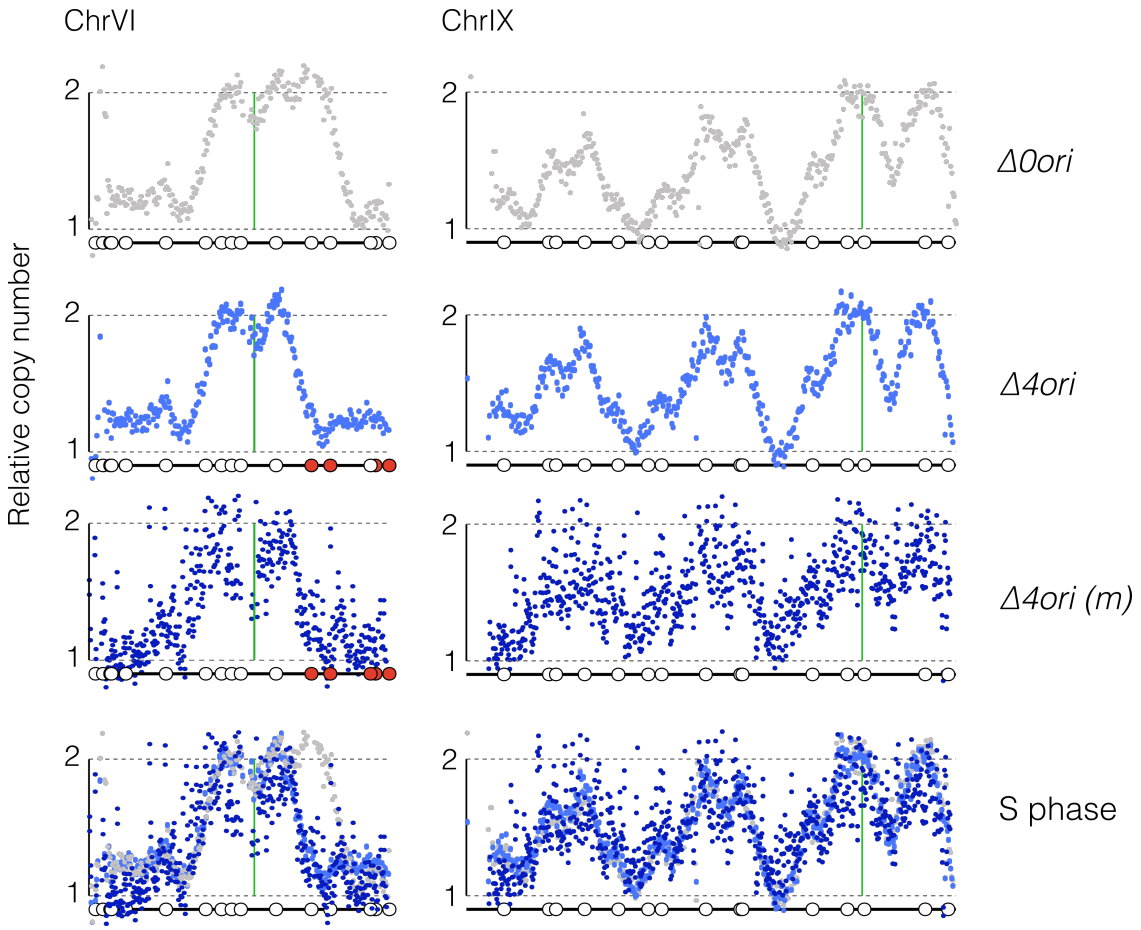


Figure 5.5 Telomeric initiation in a* Δ4ori (m) *strain is not observed by sort-seq. *Sort-seq on the* Δ4ori (m) *strain was carried out and compared to the* Δ4ori *strain. Relative copy number data for ChrVI and control ChrIX is shown. S phase* Δ4ori (m) *replicating samples were normalised to stationary phase* Δ4ori (m) *samples.* Δ4ori *data is reproduced from Figure 4.22 for comparison, and is normalised to stationary* Δ5ori *cells.*

5.3 Examining late-replicating DNA across a large replicon

Having shown that synchronous culture experiments are capable of resolving the timing of late-replicating large replicons which cannot be resolved by methods using sorted S phase cells, a synchronous culture experiment was performed in

Δ*5ori* and Δ*6ori* strains. Yeast were arrested in alpha factor and released synchronously through S phase, with samples collected for flow cytometry at regular time intervals. In order to make comparisons between yeast strains, a Gompertz function was applied to median DNA content values from flow cytometry which outputted asymmetric sigmoidal functions (Figure 5.6). From the resulting best fit functions, percentage of bulk DNA replication values were obtained for each time point.

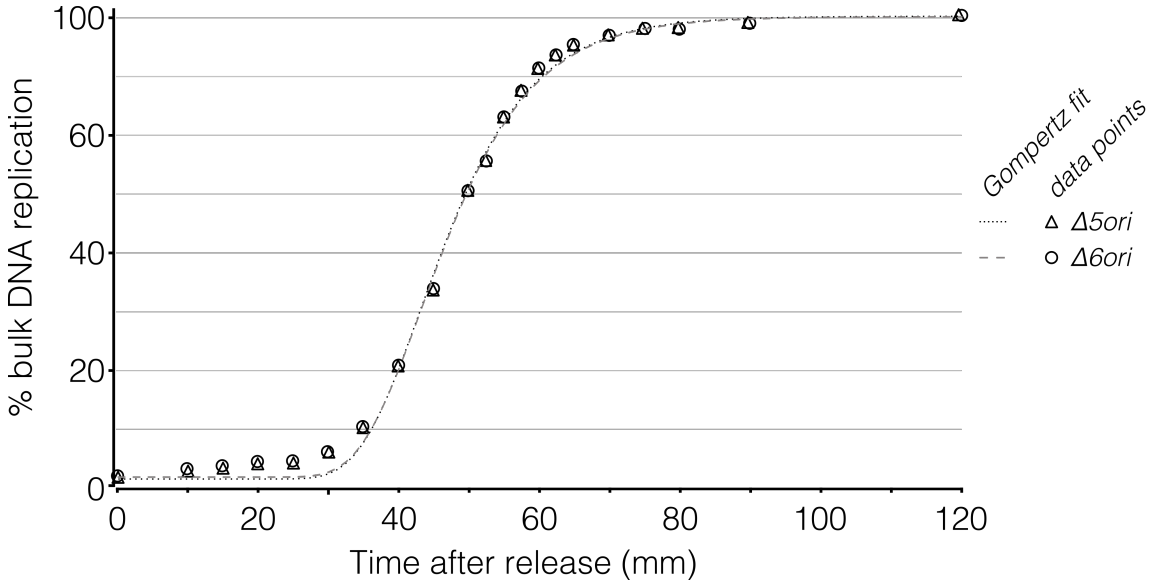


Figure 5.6 Gompertz fit applied to* Δ5ori *and* Δ6ori *cells. Δ5ori *and* Δ6ori *cells were arrested in alpha-factor and released using pronase into S phase. Cells were grown in YPAD, at 30 °C, for two hours. Flow cytometry values were calculated, and the median values inputted to an algorithm which applied a Gompertz fit. The best fit function lines are shown (spotted for* Δ5ori*, dashed for* Δ6ori*). The percentage of bulk DNA replication was outputted and plotted (triangles for* Δ5ori*, circles for* Δ6ori*). This was used to select time points with equivalent bulk DNA replication between the* Δ5ori *and* Δ6ori *strains.*

Both origin deletion strains showed similar release kinetics, with S phase lasting from approximately 30 minutes after release until 60 minutes after release (Figure 5.7). For each strain, time points with bulk DNA replication of 70 % and 95 % were selected for genomic DNA extraction and subsequent sequencing (Figure 5.7). These percentage bulk replication values were selected based on the data from Shin-ichiro Hiraga (Figure 5.4) to allow comparison of these experiments and maximise chances of observing telomeric initiation. It is important to note that measures of bulk DNA replication are not impacted by the large replicon on ChrVI. This replicon represents at most 135 kb (in the ∆*6ori* strain) which is approximately one percent of the yeast genome. Samples were normalised to genomic DNA extracted from alpha factor-arrested ∆*5ori* cells. This non-replicating sample had uniform distribution of reads per bin across the genome (data not shown).

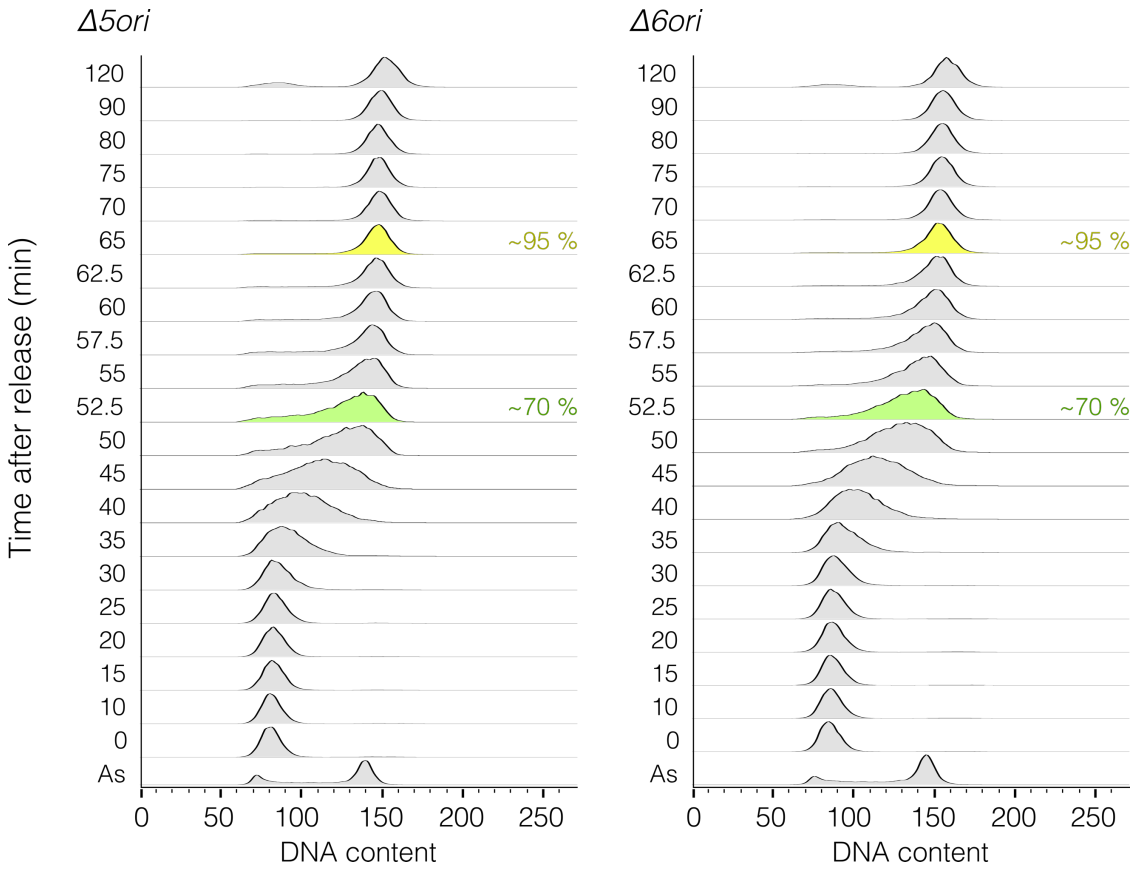


Figure 5.7* Δ5ori *and* Δ6ori *cells release synchronously from alpha-factor arrest and progress through S phase. *Δ5ori and Δ6ori cells were arrested in alpha-factor and released using pronase into S phase. Cells were grown in YPAD, at 30 °C. Flow cytometry profiles are shown for both strains at multiple time points in the two hours after release. Time points selected for DNA extraction and sequencing are highlighted in green (70 % replicated) and yellow (95 % replicated).*

At 70 % bulk DNA replication, both *Δ5ori* and *Δ6ori* strains showed an increase in relative copy number at the far right of ChrVI, with a maximum relative copy number of approximately 1.5 in the *Δ5ori* strain, and 1.25 in the *Δ6ori* strain (Figure 5.8). The inactivation of ARS606 in the Δ6ori strain was apparent, with no

peak in relative copy number at the origin, where a clear peak existed in the Δ*5ori* strain (Figure 5.8).

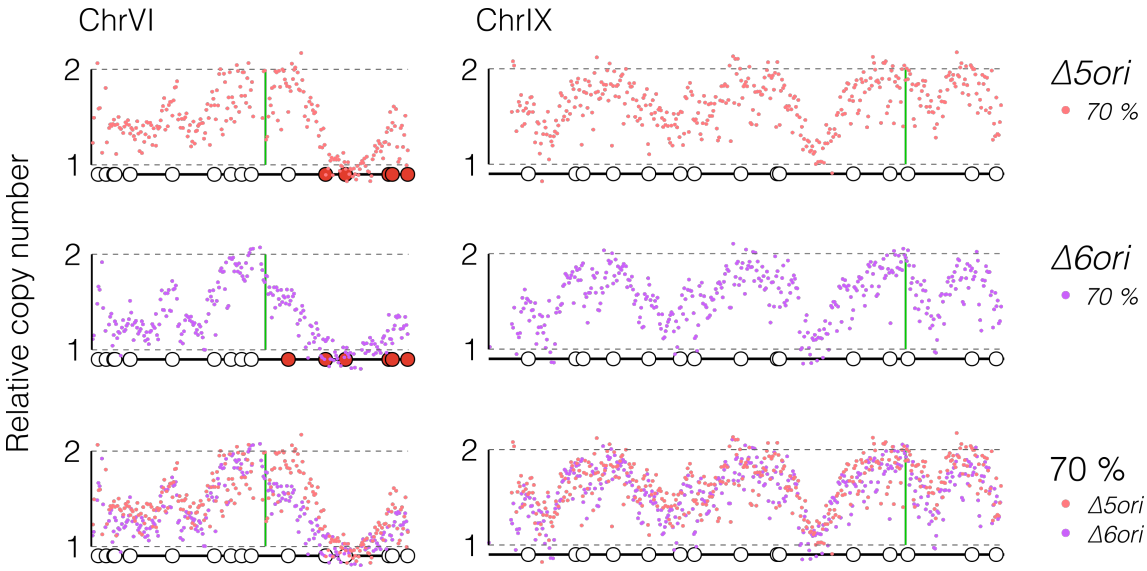


Figure 5.8 Replication of a large replicon in* Δ5ori *and* Δ6ori *strains occurs late during S phase and initiates near the telomere. *A synchronous culture experiment was carried out in* Δ5ori *(peach) and* Δ6ori *(pink) strains. Cells were grown in YPAD, at 30 °C and DNA was collected at specified time points. Relative copy number of ChrVI and control ChrIX are shown at 70 % bulk genome replicated, normalised to genomic DNA extracted from alpha factor-arrested* Δ5ori *cells.*

The 95 % bulk DNA replication data in the Δ*5ori* strain failed FASTQ quality control of DNA sequencing reads and was excluded from analysis. In the Δ*6ori* strain, the 95 % bulk DNA replication data showed an increase in relative copy number towards the far right of ChrVI (Figure 5.9). This was a higher maximum relative copy number (approximately 1.5) than the same locus at 70 % bulk

replicated (approximately 1.25) which indicates that further telomeric initiation occurred. The maximum relative copy number observed at the far right of ChrVI in the *Δ4ori (m)* strain was also 1.5, suggesting that telomeric initiation is limited to 50 % initiation.

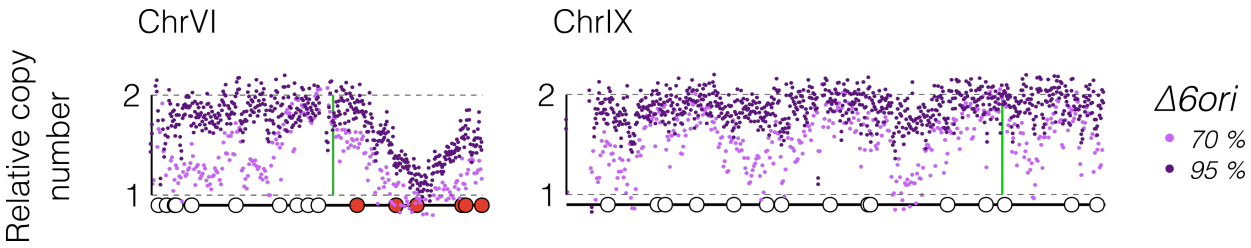


Figure 5.9 Replication forks initiating near the telomere in* Δ6ori *have limited initiation. *A synchronous culture was grown in YPAD, at 30 °C and DNA was collected at specified time points. Relative copy number of ChrVI and control ChrIX is shown at 70 % (pink) and 95 % (deep purple) bulk DNA replicated, normalised to genomic DNA extracted from alpha factor-arrested* Δ5ori *cells.*

The *Δ4ori (m)* strain at 95 % bulk DNA replication had a relative copy number of 2.0 between approximately 1 kb and 200 kb (Figure 5.4B). However, the large replicon in the *Δ6ori* strain at 95 % bulk DNA replication had relative copy number of 2.0 between approximately 1 kb and 160 kb (Figure 5.9). This indicates that a delay to replication of the right arm of ChrVI results from the further inactivation of the early-firing origin *ARS606*. Between 70 % and 95 % bulk DNA replication the entire left arm of ChrVI has replicated, including loci at a low relative copy number at 70 % bulk DNA replication. The increase in relative copy number on the right arm of ChrVI is more modest, indicating that replication completion is dependent on incoming replication forks from*ARS605*. This is

consistent with the limited replication 'competence' of telomeric initiation discussed above.

5.4 Investigating polymerase-usage normalisation methods

5.4.1 Methods for normalising ribonucleotide incorporation data

HydEn-seq and Pu-seq both use polymerase mutant strains which incorporate ribonucleotides at elevated rates, in the context of *RNH201* deletion, in order to investigate strand bias of polymerases (Figure 5.10A) (Clausen *et al.* 2015; Keszthelyi *et al.* 2015). However, the method used to normalise the raw ribonucleotide incorporation data differs between the two studies.

Clausen *et al*. investigated the strand bias of Polε, Polδ and Polα independently, normalising the incorporation of ribonucleotides on the Watson strand of DNA against the ribonucleotides incorporated on both strands in that locus (Figure 5.10B, Method 1). They further removed background levels of ribonucleotides which were not inserted by the promiscuous polymerase by subtracting the ribonucleotides signal produced by a ∆*rnh201* yeast strain (Clausen et al. 2015). In contrast, Keszthelyi et al. investigated the relative strand bias of Polε and Polδ by normalising the incorporation of ribonucleotides on the Watson strand of DNA in the Polε strain against the sum of the ribonucleotides incorporated on the Watson strand of DNA in the Polε and Polδ strains, and vice versa (Figure 5.10B, Method 2) (Keszthelyi et al. 2015). A mixture of these methods was used to investigate polymerase-usage of Polε and Polδ across the large replicon in origin deletion strains. Generally, these two normalisation methods give equivalent results, which is what is anticipated based on the division of labour at the replication fork (Johansson & Dixon 2013).

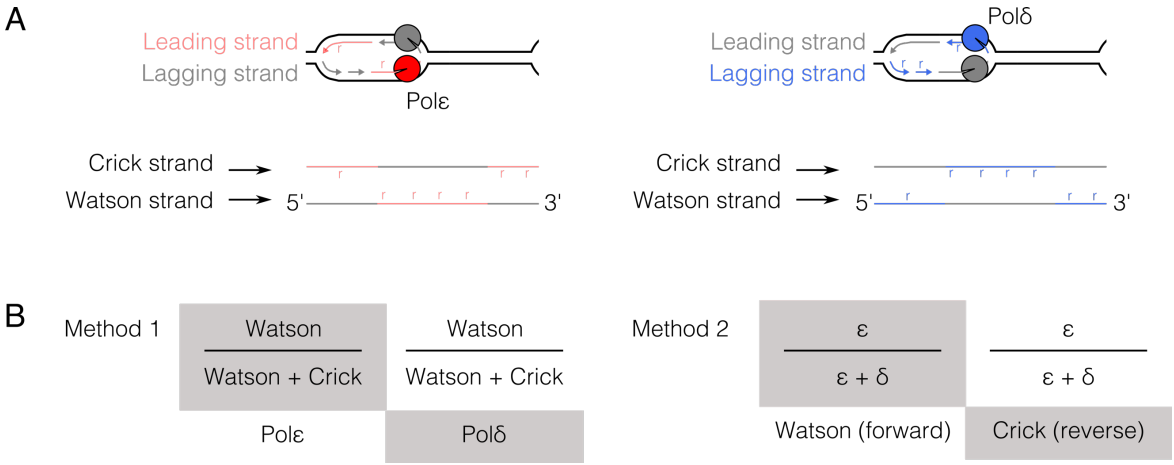


Figure 5.10 Two methods to normalise ribonucleotide incorporation data. *(A) Mutant forms of Polε (*pol2-M644G*) or Polδ (*pol3-L612G*) incorporate ribonucleotides at elevated rates predominantly into the leading or lagging strands respectively. (B) Method 1 normalises data from the Watson strand in one yeast strain against data from both strands of the same strain. Method 2 requires complementary yeast strains to normalise the Watson strand data from one yeast strand (e.g.* pol2-M644G*) against the Watson strand data from both strains (e.g.* pol2-M644G *and* pol3-L612G*). Methods 1 and 2 are used in (Clausen et al. 2015) and (Keszthelyi et al. 2015) respectively.*

5.4.2 Pol2-dependent replication in Δ4ori *cells*

Initially, a single yeast strain was investigated to examine a 100 kb replicon on ChrVI: *Δ4ori pol2-M644G*, *Δrnh201*. An asynchronous population of cells was grown in YPAD at 30 °C to an OD_{600} of 0.8 (1 x 10^7 cells/ml). Cells were harvested and DNA was extracted. Extracted DNA was alkali-treated, breaking the covalent bonds 3' of ribonucleotide incorporation. The resulting ssDNA had a second

strand synthesised, using dUTP instead of dTTP and was used as template in a modified Next-Seq library preparation. During this process adaptors were added followed by USER enzyme digestion, which also degraded the uracil-containing second strand. The result was 5' adaptor ends which specifically denoted the site 3' of ribonucleotide incorporation. After mapping, the sites 5' of 5' ends were recorded as ribonucleotides.

The ∆*4ori pol2-M644G* ∆*rnh201* data were normalised using a simplified version of the Clausen *et al*. process, Method 1 in Figure 5.10B (Clausen *et al.* 2015). Ribonucleotide incorporation by Polε into the Watson (forward) strand is plotted in Figure 5.11A. A sharp increase of ribonucleotide incorporation (5' to 3') by Polε into the Watson indicates on origin site, while a decrease in signal indicates termination zones. The extent of the increase in ribonucleotide signal represents the efficiency of the origin – highly efficient origins are shown by a transition from very low signal to very high signal over their loci. This interpretation of ribonucleotide signal assumes that Polε synthesises the leading strand, and Polδ synthesises the lagging strand (Johansson & Dixon 2013).

The ∆*4ori pol2-M644G* ∆*rnh201* strain showed an increase in ribonucleotide incorporation at *ARS606* equivalent to that seen in the ∆*0ori pol2-M644G* ∆*rnh201* strain (Figure 5.11). Ribonucleotide incorporation remained high over *ars607*, indicating the origin was inactive. However, just prior to *ars608* the ribonucleotide incorporation signal on the Watson strand began to reduce (Figure 5.11A). It then fluctuated but remained between approximately 25 % and 75 % across the large replicon. This suggests, contrary to expectation (Figure 5.3B) that Polε was not fully responsible for replicating the Watson strand in the ∆*4ori*

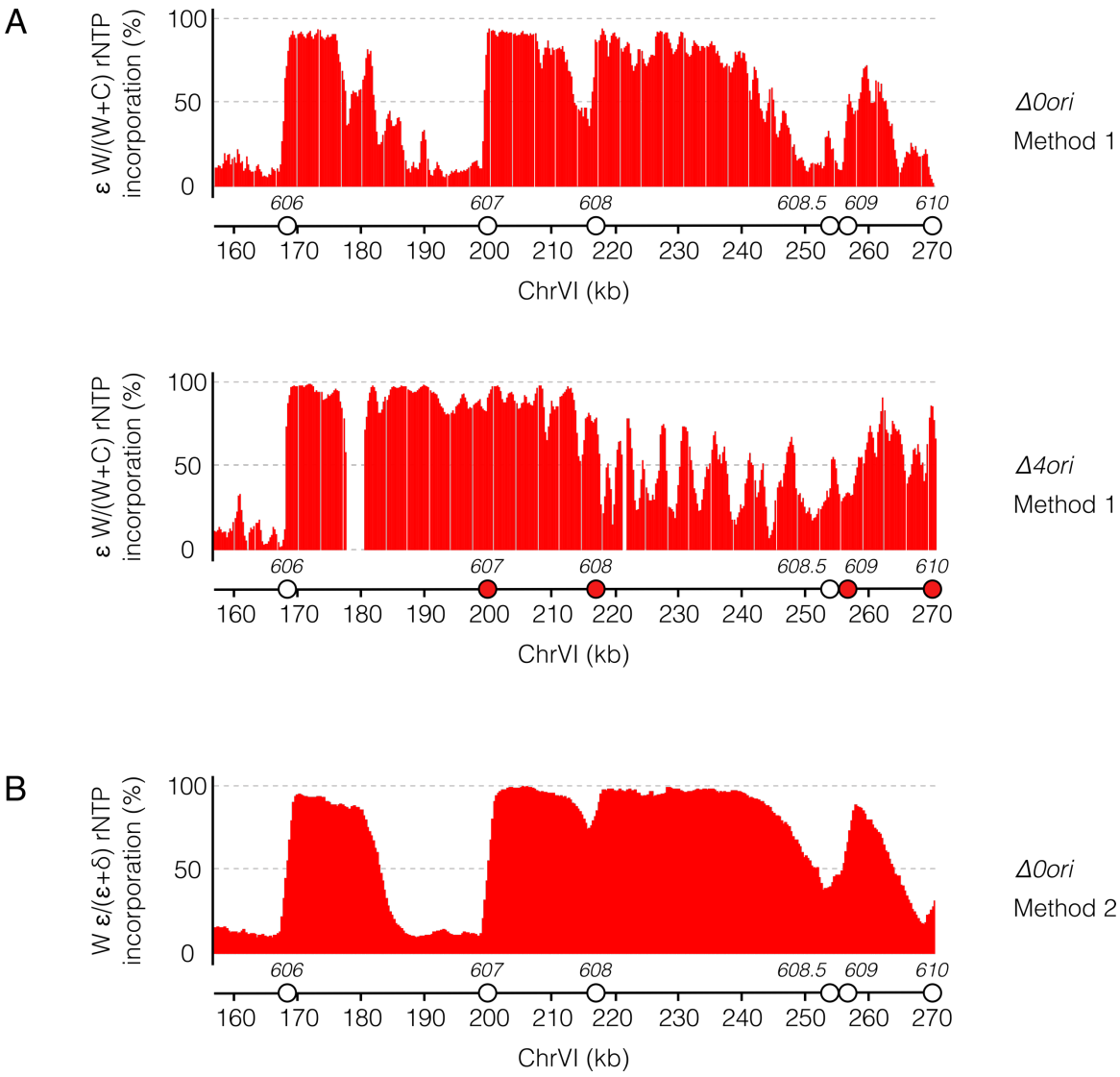


Figure 5.11 Leading strand synthesis in* Δ4ori *cells is not exclusively due to Pol-2 replication initiating at ARS606. *(A) Asynchronous* Δ4ori pol2-M644G Δrnh201 *cells were harvested and used for Pu-seq. Sequencing data were normalised using Method 1 (Figure 5.10B).* Δ0ori pol2-M644G Δrnh201 *data from Katie Ptasińska (unpublished) were re-analysed using the same normalisation method for comparison. (B)* Δ0ori pol2-M644G Δrnh201 *data from Katie Ptasińska (unpublished) normalised using a complementary* Δ0ori pol3-L612G Δrnh201 *strain - Method 2 in Figure 5.10B.*

pol2-M644G Δ*rnh201* strain across the large replicon (compare Figure 5.3B to Figure 5.11A, Δ*4ori*).

The fluctuations observed in ribonucleotide incorporation across the large replicon could reflect rapid changes in polymerase usage of Polε in the Δ*4ori pol2-M644G* Δ*rnh201* strain. However, when the Δ*0ori pol2-M644G* Δ*rnh201* data from Katie Ptasińska were normalised in the same manner using Method 1 (Figure 5.11A) they showed increased fluctuation relative to the same data normalised to a complementary Polδ strain (compare Figure 5.11A, Δ*0ori* to Figure 5.11B, Δ*0ori*). Despite this, it is worth noting that the extent of fluctuation observed in the Δ*4ori pol2-M644G* Δ*rnh201* strain was far greater than the fluctuation in the Δ*0ori pol2-M644G* Δ*rnh201* strain (Figure 5.11A). Therefore, it was concluded it was preferable to investigate relative polymerase usage by normalising between pairs of complementary Polε and Polδ strains following the published Pu-seq protocol (Keszthelyi *et al.* 2015).

5.5 Polymerase-usage sequencing of Δ*5ori* and Δ*6ori* cells

S. cerevisiae strains were created which were *pol2-M644G* Δ*rnh201* or *pol3-L612G* Δ*rnh201* with either Δ*0ori,* Δ*5ori* or Δ*6ori* origin deletions (henceforth complementary pairs are referred to as Δ*0ori,* Δ*5ori* or Δ*6ori* Pu-seq strains). Each yeast strain was grown, harvested and had DNA extracted and sequenced as described in 5.3.2. Ribonucleotide signal was normalised from Polε strains using complementary Polδ strains, as in Figure 5.10B (Method 2). The Δ*0ori* Pu-seq strains were found to have a ribonucleotide signal profile similar to the *S. cerevisiae* Δ*0ori* Pu-seq data from Katie Ptasińska (compare Figure 5.11 B to

Figure 5.12 ∆*0ori*, forward strand). This showed the Pu-seq technique to be reproducible in independent strains.

For the following discussion, unless otherwise stated comments on Pu-seq data refer to the forward strand. The ∆*5ori* Pu-seq strains showed an increase in ribonucleotide signal at *ARS605* and *ARS606* comparable to the ∆*0ori* Pu-seq strains (Figure 5.12). In the ∆*0ori* Pu-seq strains the activity of *ARS608* was visible as a decrease in ribonucleotide signal just prior to the origin site (a sharp increase), where leftwards moving forks incorporated ribonucleotides using Polδ. No decrease in ribonucleotide signal followed by sharp increase was seen prior to *ars607* or *ars608* in the ∆*5ori* Pu-seq strains, indicating Polε was responsible for all synthesis of the Watson strand in this region of the replicon (Figure 5.12). Instead, the ribonucleotide signal began to decrease just prior to *ars608* and continued to decrease on a gradual incline until *ars609*.

The gradual decrease in ribonucleotide signal seen in the ∆*5ori* Pu-seq strains resembles a termination zone. In termination zones, oncoming forks meet stochastically, producing a lesser gradient of change in signal (e.g. in Figure 5.12 ∆*0ori* Pu-seq strains, between *ARS608* and *ARS608.5*) relative to the sharp increases in signal observed at origin sites. The same gradual decline in ribonucleotide signal was observed in the ∆*6ori* Pu-seq strains; this decline began just after *ars606* and the level of ribonucleotide signal at each point along the replicon was reduced relative to that seen in the ∆*5ori* Pu-seq strains (Figure 5.12). Conversely, the ribonucleotide signal on the reverse strand of the ∆*6ori* Pu-seq strains was also higher at each location along the replicon than in the ∆*5ori* Pu-seq strains (Figure 5.12). These data supports the hypothesis that the gradual decline observed represents a termination zone, where forks incoming from the

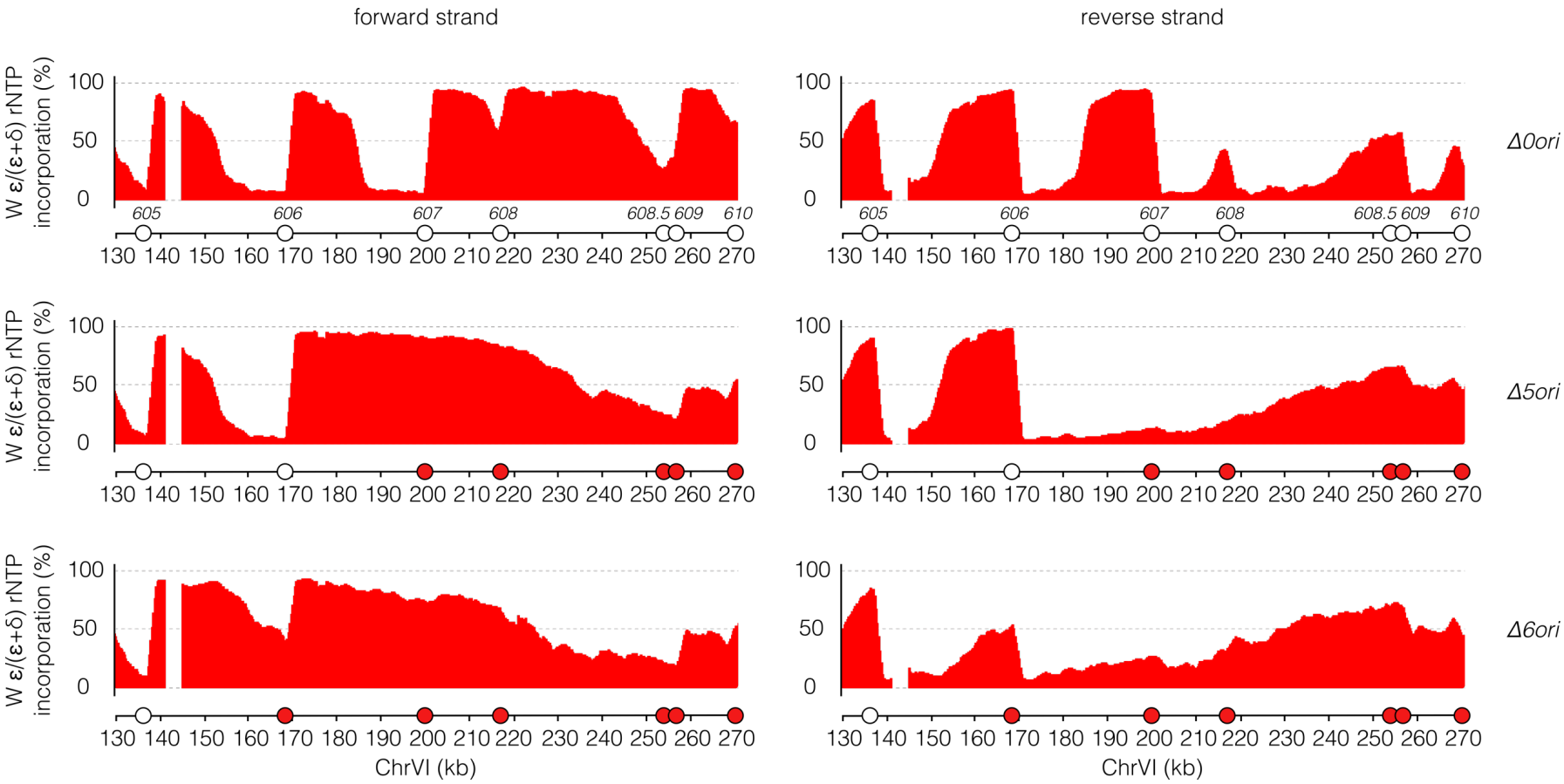


Figure 5.12 Replication of the large replicon in* Δ5ori *and* Δ6ori *yeast is due in part to replication initiating close to telomeres. *Asynchronous* pol2-M644G Δrnh201 *cells (*Δ0ori, Δ5ori *or* Δ6ori*) were harvested and used for Pu-seq. Sequencing data was normalised using complementary* pol3-L612G Δrnh201 *cells (*Δ0ori, Δ5ori *or* Δ6ori*) strains - Method 2 (Figure 5.10B).*

far right of ChrVI encounter replication forks originating at *ARS606* in the ∆*5ori* Pu-seq strains and at *ARS605* in the ∆*6ori* Pu-seq strains. Since forks, on average, have 35 kb further to travel across the replicon in the ∆*6ori* Pu-seq strains than in the ∆*5ori* Pu-seq strains there is a leftwards shift in the termination zone, observed in the Pu-seq data here.

Using the complementary Polδ strain to normalise Pu-seq data means that in areas with low Polε and Polδ usage (e.g. in hard-to-replicate areas where another polymerase takes over synthesis) the normalised data will not show this decrease in overall canonical polymerase usage. Therefore, the number of raw reads mapping across the genome, both prior to and after ribonucleotide assignment, was plotted (Figure 5.13). Neither genome-wide coverage of reads (Figure 5.13A) nor mapping of ribonucleotide sites (Figure 5.13B) decreased at the far right of ChrVI relative to the *ARS605* active origin or the rest of the genome (data not shown). In the ribonucleotide mapping data, an increase in ribonucleotide incorporation by *pol3-L612G* was seen on the forward strand of ∆*5ori* and ∆*6ori* Pu-seq strains (Figure 5.13B). The inverse was observed on the reverse strand, where an increase in ribonucleotide incorporation by *pol2-M644G* was seen in ∆*5ori* and ∆*6ori* Pu-seq strains (data not shown). Therefore, the changes in Polε and Polδ usage seen in ∆*5ori* and ∆*6ori* Pu-seq strains would seem to represent a genuine increased usage of Polδ on the Watson strand and increased usage of Polε on the Crick strand, consistent with replication forks traversing inwards from the far right of ChrVI (Figure 5.12).

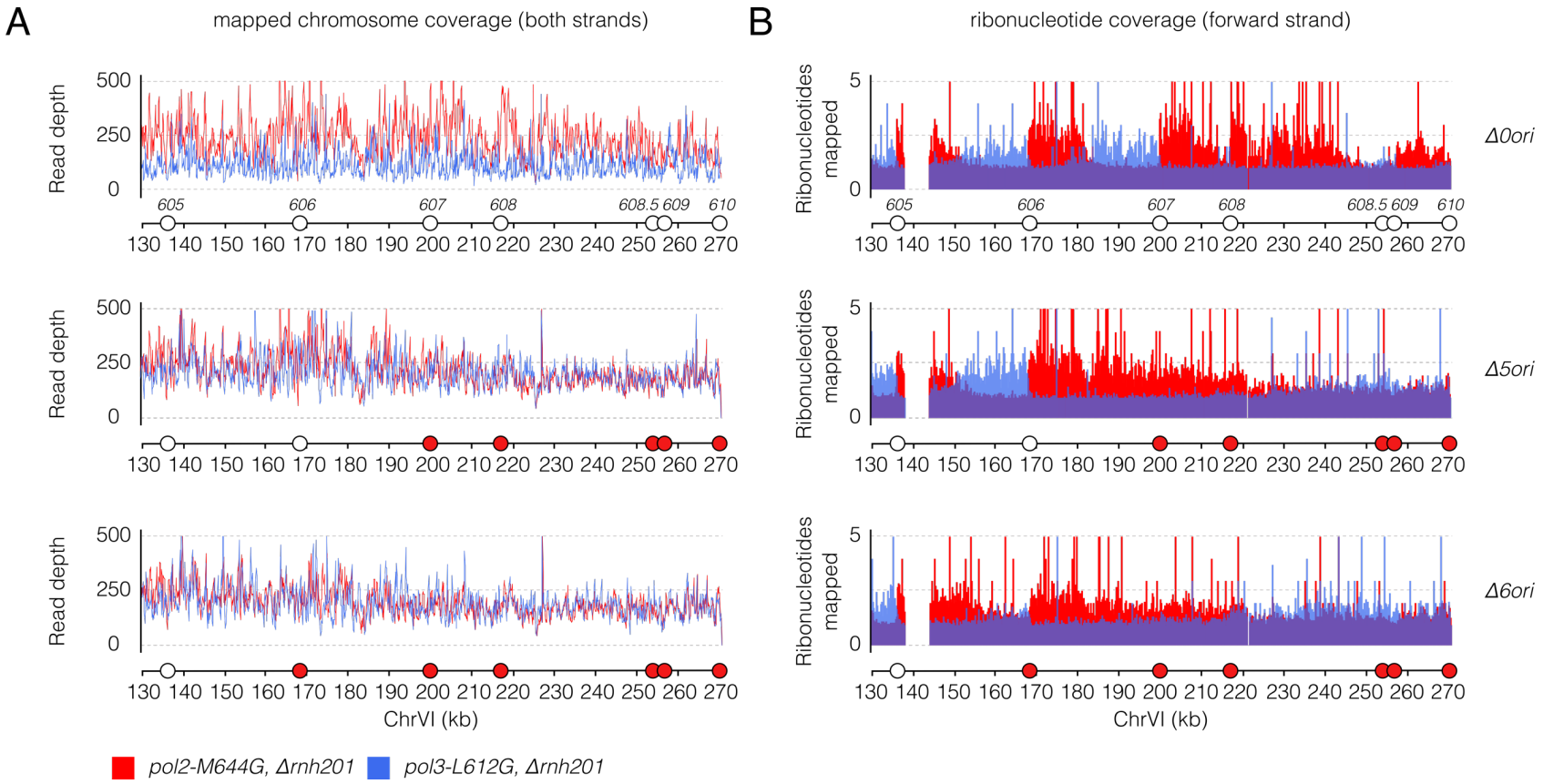


Figure 5.13 Pol2 and Pol3 are responsible for the majority of replication across the large replicon in* Δ5ori *and* Δ6ori *yeast. *Sequencing coverage is shown for* pol2-M644G Δrnh201 *cells (*Δ0ori, Δ5ori *or* Δ6ori*) in red,* pol3-L612G Δrnh201 *cells (*Δ0ori, Δ5ori *or* Δ6ori*) in blue. (A) the genome coverage on both strands and (B) the number of ribonucleotides mapped to the right arm of ChrVI are plotted. Data is smoothed into 300 bp bins.*

A second factor that could impact on normalising ribonucleotide data is a difference in propensity for the mutant polymerases to incorporate specific ribonucleotide bases. The base for each ribonucleotide site was called using a SAMTOOLS script and the percentage of each ribonucleotide incorporated by the mutant *pol2-M644G* or *pol3-L612G* calculated (Figure 5.14). It was found that across the genome both mutant polymerases had similar predispositions to incorporate each ribonucleotide base, following the preference order: rC (~32 %) > rG (~30 %) > rA (~ 23 %) > rT (~ 15 %). This indicated approximately 64 % of ribonucleotides incorporated were G or C bases, which was a surprise given the *S. cerevisiae* genome is 62 % AT bases. However, this finding was supported by the same preference hierarchy being found in the 10 % most strand-biased areas of the genome by Clausen *et al* (Clausen *et al.* 2015).

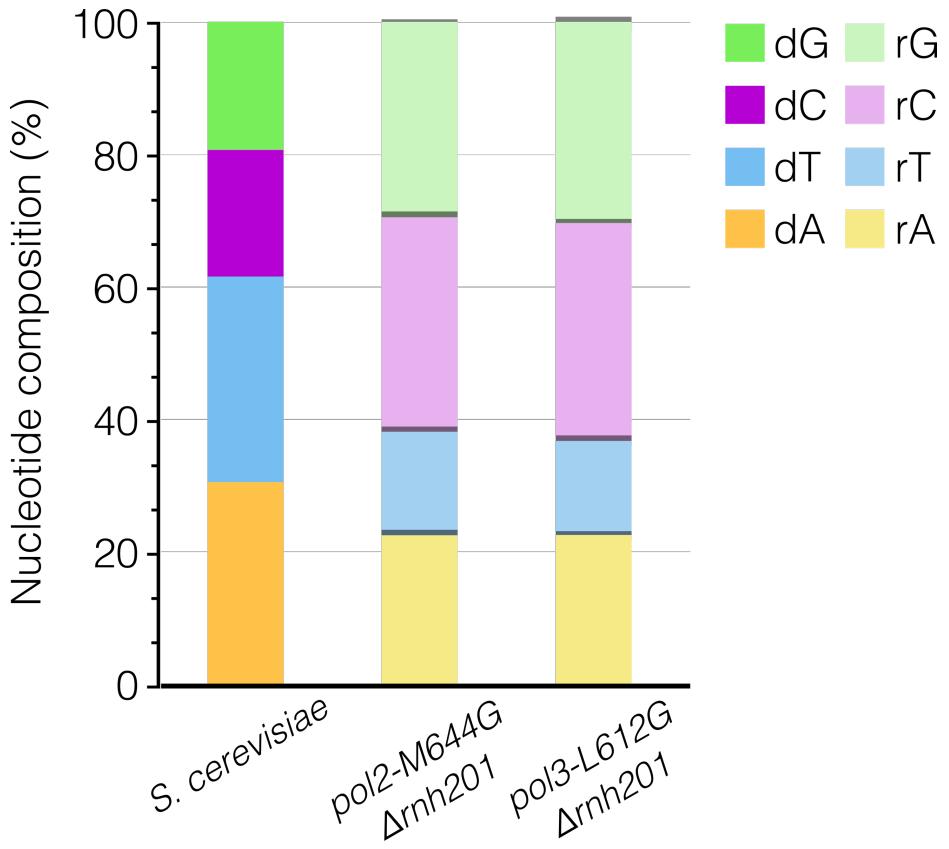


Figure 5.14* pol2-M644G *and* pol3-L612G *incorporate ribonucleotides preferentially at G and C positions. *Genome wide GC content for* S. cerevisiae *reference genome (SacCer3) is 38 % (UCSC). The genomic percentage incorporation of each ribonucleotide by* pol2-M644G *and* pol3-L612G *was calculated for* Δ0ori, Δ5ori *and* Δ6ori *strains and the mean was plotted. Positive standard deviation is shown in grey above each clustered histogram section.*

5.6 'Inactive' origins *ars606* and *ars609* appear capable of inefficient initiation

In the Pu-seq data above, sharp increases in ribonucleotide signal were seen in locations close to *ars606* in the Δ*6ori* Pu-seq strain and *ars608.5*/*ars609* in the Δ*5ori* and Δ*6ori* Pu-seq strains (Figure 5.12). When magnified in the IGV viewer,

it was found that there was an increase in ribonucleotide signal in the Δ*6ori* strain at *ars606*. This aligned to the increase in ribonucleotide signal representing *ARS606* in the Δ*0ori* Pu-seq strain (Figure 5.15). This appeared to suggest that there was some residual activity from the 'inactive' *ars606* in the Δ*6ori* Pu-seq strain.

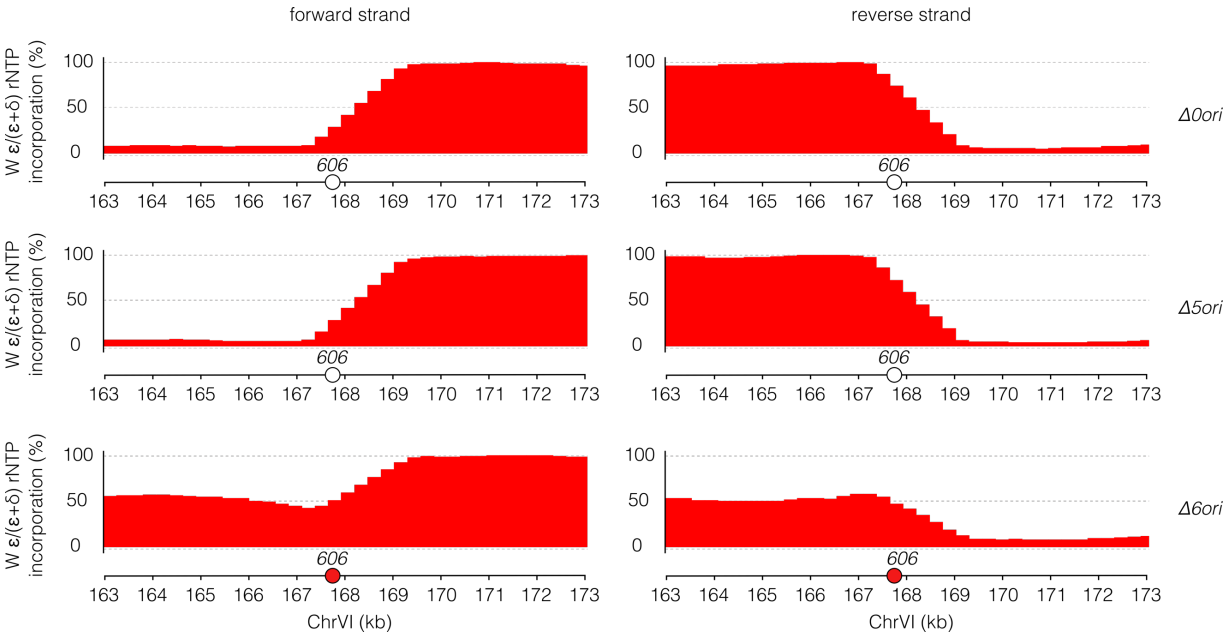


***Figure 5.15 'Inactive'* ars606 *is able to initiate replication at a reduced efficiency to* ARS606.** *A higher resolution region of Figure 5.12 is plotted around* ARS606/ars606.

Similarly, the *ars608.5/ars609* region was magnified in IGV. This showed that the increase in ribonucleotide signal observed in the Δ*5ori* and Δ*6ori* Pu-seq strains around *ars608.5/ars609* aligned to *ARS609* in the Δ*0ori* Pu-seq strain (Figure 5.16). No corresponding increase in ribonucleotide signal at *ARS608.5* in the Δ*0ori* Pu-seq strain was observed over *ars608.5* in either the Δ*5ori* or the

Δ*6ori* Pu-seq strains showing that the inactivation of this potential origin site was effective.

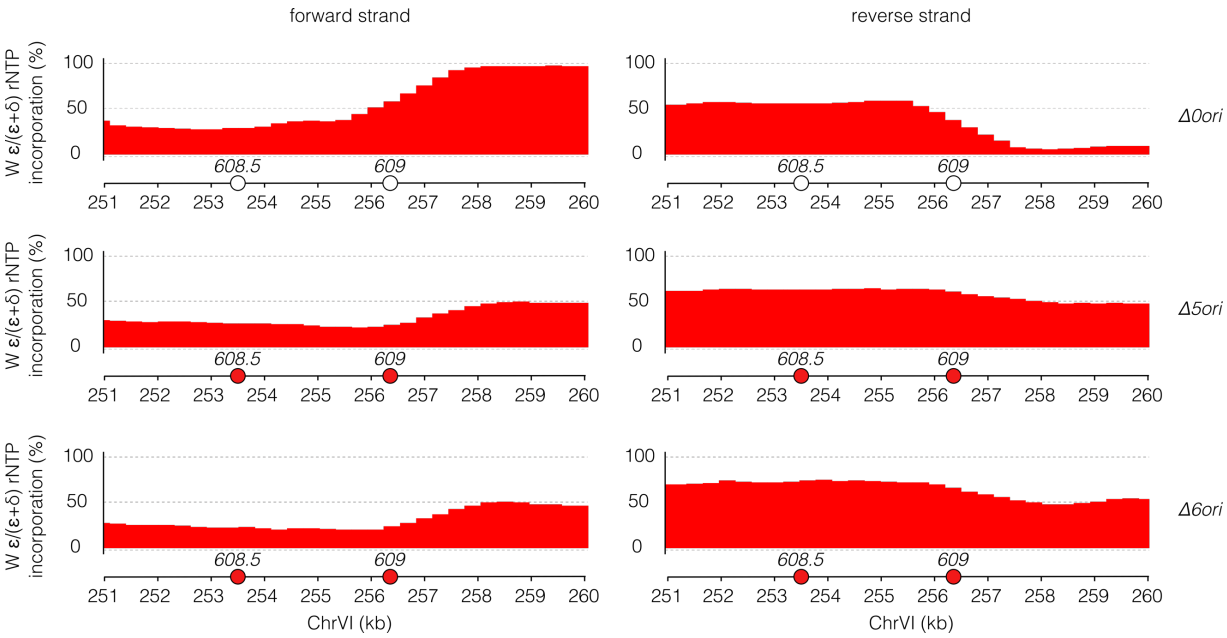


***Figure 5.16 'Inactive'* ars609 *is able to initiate replication at a reduced efficiency to* ARS609.** *An amplified region of Figure 5.12 is plotted around* ARS609/ars609.

The increase in ribonucleotide signal observed in the Δ*6ori* Pu-seq strain at *ars606* contrasted with published synchronous culture experiment data where the same *ars606* origin inactivation sequence led to an almost complete removal of origin activity at 15 and 20 minutes into S phase (Hawkins *et al.* 2013b). It also contrasted with the synchronous culture (Figure 5.8) and sort-seq (Figure 4.23) experiments performed in this thesis, neither of which saw origin activity from *ars606*. It was therefore necessary to confirm that the *ars606* origin site had not mutated in the Δ*6ori* Pu-seq strain. Sequencing data confirmed that both the

ars606 and *ars609* sites from the Δ*6ori* Pu-seq strain were mutated relative to the *ARS606* and *ARS609* sequences, and were not similar to the ACS (Figure 5.17). These sequences confirmed the expected origin inactivation mutations were present in the Δ*6ori* Pu-seq strain (Table 4.1).

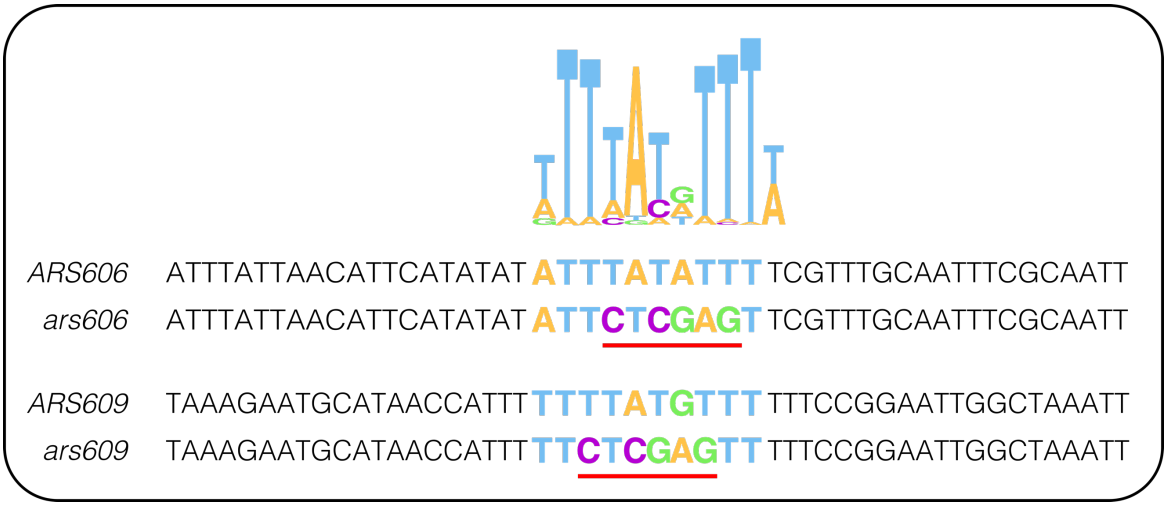


Figure 5.17 Sequencing data confirm the modification of* ars606 *and* ars609 *ACS sites. *The 11bp consensus sequence is shown as a likelihood plot (modified from (Descorps-Declère et al. 2015)). ACS sequences as annotated in UCSC are coloured to match the consensus.* ARS606 *and* ARS609 *sequences were taken from USCS with the surrounding 20 bp of DNA;* ars606 *and* ars609 *sequences were taken from sequencing data of the* Δ6ori *strain. The replacement sequences introduced an* XhoI *restriction enzyme site, underlined in red.*

The published inactivation of *ars606* was unable to support colony growth in a plasmid assay (Hawkins *et al.* 2013b). A plasmid assay was performed to confirm whether *ars606* was capable of supporting plasmid replication (Figure 4.10). Yeast were transformed with plasmids with either *ARS606* or *ars606* using

at least 600 bp of surrounding DNA, then plated onto selective medium. Plasmids containing wild-type *ARS607* or mutant *ars607* were used as controls. After two days, the published result was confirmed; multiple large colonies grew in a strain with a *ARS606* plasmid whereas no colonies were seen when as *ars606* plasmid was used instead (Figure 5.18).

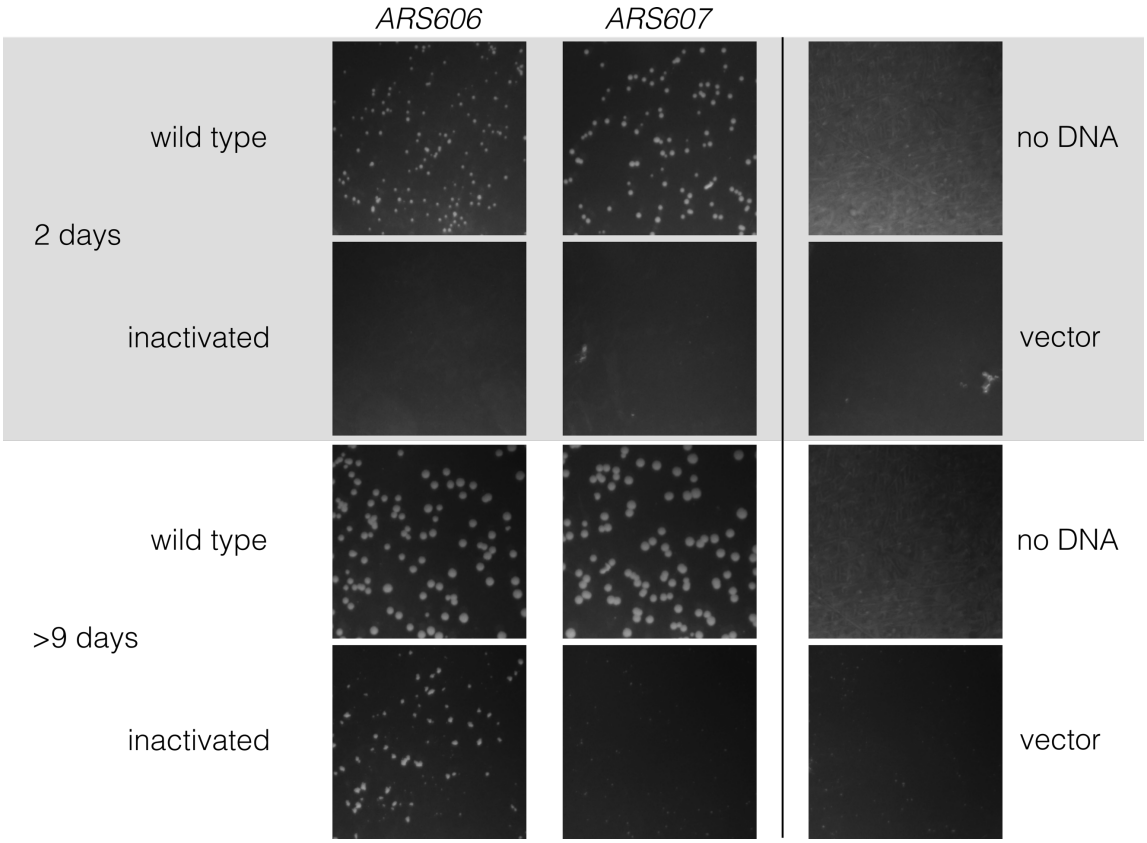


Figure 5.18* ars606 *supports irregular colony growth after extended incubation. *A plasmid assay was carried out, as described in Figure 4.10. Plates were incubated at 30 °C for 9 days, then at room temperature for a further 2 days (where growth was minimal). Plates were imaged after 2 days at 30 °C and after 9 days at 30 °C and 2 days at room temperature. Vector only and no DNA control transformations are also shown.*

However, after an extended incubation (for 9 days at 30 °C and 2 days at room temperature) irregular colonies were seen in the yeast strain with an *ars606* plasmid (Figure 5.18). This was in contrast to the yeast containing an *ars607* plasmid, where no colonies were viewed after either 2 days or the extended incubation (Figure 5.18). This suggested that there was reduced origin activity from the mutant origin, therefore, the origin was not completely inactive. This may be due to a secondary inefficient ORC binding site that remains in the mutant origin. Conversely, the complete inactivation of *ARS607* by the *ars607* sequence was confirmed.

5.7 Modelling modes of ChrVI right replication

Replication timing is reproducible between cell cycles. Previous work by the Nieduszynski lab showed that mathematical modelling can simulate replication timing (Hawkins *et al.* 2013b), where each replication origin is denoted by two features – its mean replication timing and its efficiency (the percentage of cells in which the origin fires). Recently, Michael Boemo (Nieduszynski laboratory) created a computational language to model biological processes, such as tumour-associated macrophages (Boemo *et al.* 2019; Boemo & Byrne 2019). Using modelling, it is possible to predict the impact of changes, such as introducing a large replicon on replication timing.

In this chapter, a DNA replication-specific adaptation of Michael Boemo's computational language (Boemo *et al.* 2019) was applied to model the relative usage of Polε and Polδ when replicating DNA. The usage of Polε on the Watson strand of DNA relative to total polymerase usage on the Watson strand was plotted, which mimicked the output of Pu-seq data. Values for each origin were

inputted from published mean replication timing and efficiency data (Hawkins *et al.* 2013b). The resulting model was trained on wild-type Δ*0ori* Pu-seq strain data from Katie Ptasińska. New mean replication timing and efficiency values for each origin, which closely simulated both wild-type Pu-seq data and wild-type DNA replication timing data, were generated by this training process.

A comparison of wild-type Δ*0ori* Pu-seq strain data and plotted Pu-seq simulations from the final model is shown in Figure 5.19A. A large replicon simulating the Δ*5ori* Pu-seq strain was introduced to the model by removing the nodes representing *ARS607* – *ARS610*. The model predicted this would lead to high usage of Polε across the replicon (Figure 5.19B), as was predicted in Figure 5.3B. It has been previously shown in this chapter that this prediction was incorrect; a form of replication initiation on the far right of ChrVI was supported by both synchronous culture experiments and Pu-seq data (Figure 5.8, 5.11). Alternative replication initiation was hypothesised to explain the discrepancy between the Δ*5ori* Pu-seq data and the Δ*5ori* Pu-seq model (Figure 5.19C).

Two modes of alternative replication were investigated by modelling. The first assumed non-specific initiation could occur anywhere throughout the replicon in the absence of known active origins. *In vitro* studies have shown that DNA replication can initiate on DNA in the absence of active origins (On *et al.* 2014). Furthermore, one species of archaea has been shown to replicate its circular chromosome more rapidly in the absence of any active origins, with no discrete sites of replication (Hawkins *et al.* 2013a). Single molecule sequencing has been used recently to show that replication initiation in yeast can occur at a low frequency at many previously unknown genomic loci (Müller *et al.* 2019). However, modelling of a Δ*5ori* Pu-seq strain with non-specific initiation over the

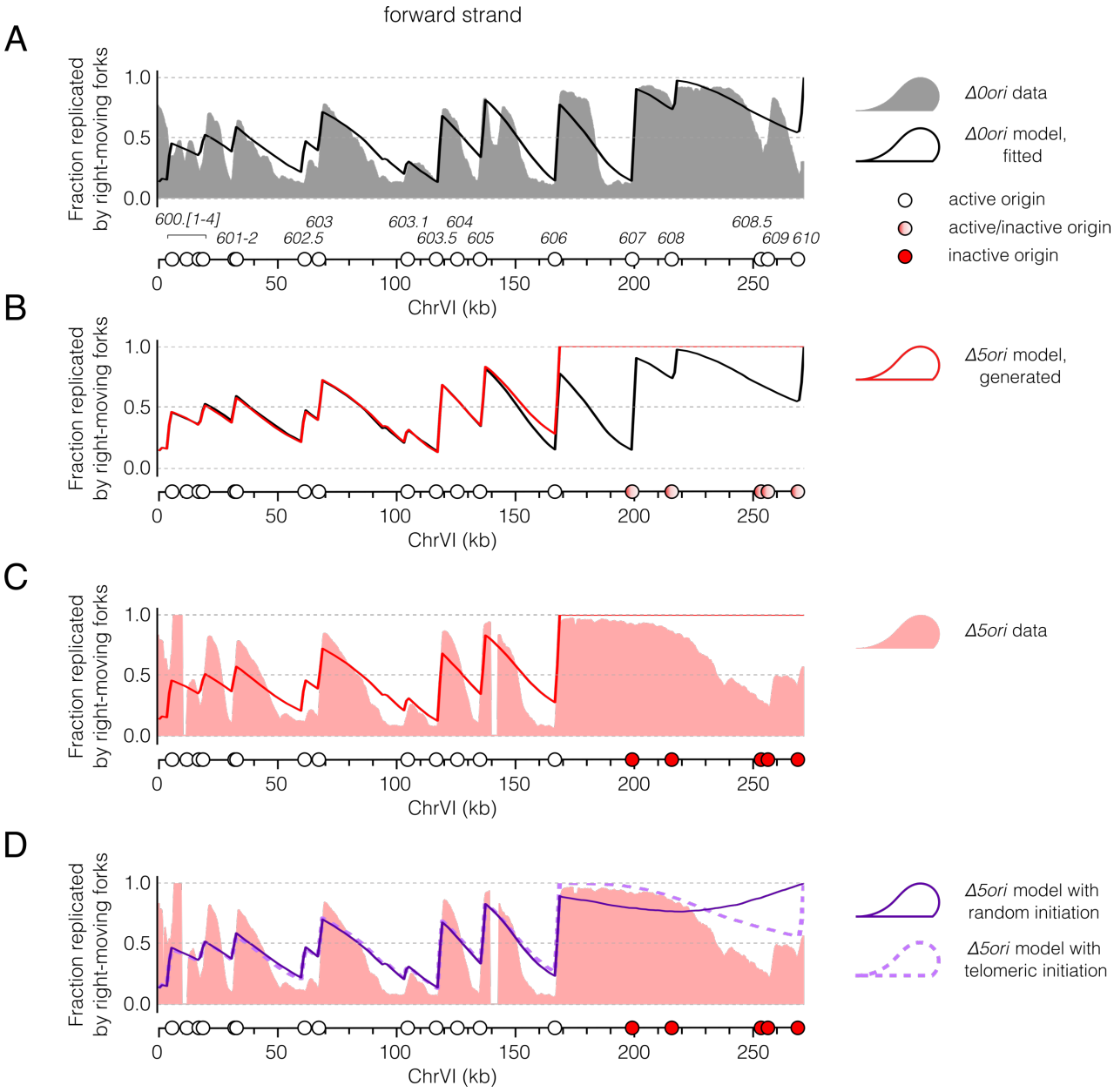


Figure 5.19 Modelling of DNA replication indicates a telomeric initiation site is responsible for leftward moving replication across the large replicon in* Δ5ori *yeast. *A mathematical modelling language for biological processes was developed by Michael Boemo* (Boemo *et al.* 2019)*. Origin parameters were inputted from (Hawkins et al. 2013b). The resulting model was (A) fitted to* Δ0ori *Pu-seq data from Katie Ptasińska and (B) subsequently used to predict* Δ5ori *Pu-seq data where all replication occurs from known origins. This model was (C)*

compared to the Δ5ori *Pu-seq data. (D) Two alternative non-canonical replication forms in the* Δ5ori *strain were modelled and compared to the* Δ5ori *Pu-seq data. Random initiation (solid purple) was modelled as a stochastic ability for replication to occur anywhere within the large replicon. Telomeric initiation (purple dashed) was modelled as a single initiation site at 270 kb.*

entirety of the replicon region did not result in a profile similar to the Δ*5ori* Pu-seq strain data (Figure 5.19D). A range of initiation frequencies were tried (data not shown), however, all simulations resulted in 100 % of replication of the Watson strand at the far right of ChrVI being performed by Polε.

The second model of alternative replication introduced a telomeric site of initiation at the end of ChrVI (270 kb). This resulted in a reduction of Polε signal across the replicon site from *ARS606* which aligned closely to the reduction in ribonucleotide signal from *ARS606* observed in the Δ*5ori* Pu-seq strain data (Figure 5.19D). A sharp increase of signal was observed in the model immediately prior to the end of ChrVI (Figure 5.19D). However, this increase was also observed in the wild-type simulated data (Figure 5.19A, B), suggesting it was a result of the modelling algorithm. Varying the mean replication time or efficiency of the telomeric initiation site changed the magnitude of reduction in Polε signal across the replicon site, nonetheless, the overall trend remained the same (data not shown). These data support a hypothesis in which replication is initiated at the telomere with oncoming replication forks stochastically terminating as they encounter replication forks initiating at *ARS606*.

5.8 Summary

This chapter aimed to build upon the findings of chapter 4 and further investigate consequences of inactivated origins in *S. cerevisiae*. It was observed in replication timing data collected by Shin-ichiro Hiraga that replication initiating from the far right of ChrVI occurred in a *S. cerevisiae* strain with a large replicon of 100 kb (Figure 5.4, unpublished). This chapter showed that this replication occurred after mid-S phase, and could not be resolved using sort-seq (Figure 5.5). Therefore, ∆*5ori* and ∆*6ori* strains examined by sort-seq in Chapter 4 were re-examined by synchronous culture experiments. It was found that ∆*5ori* and ∆*6ori* strains progressed at the same rate through S phase after alpha factor release (Figure 5.7) and replication initiating from the far right of ChrVI (termed telomeric initiation) was also apparent in these strains (Figure 5.8).

A hypothesis of telomeric initiation was supported by polymerase-usage data. Both the ∆*5ori* and ∆*6ori* Pu-seq strains showed a reduction in ribonucleotide signal due to Polε on the Watson (forward) strand and an increase in ribonucleotide signal due to Polε on the Crick (reverse) strand (Figure 5.12). This indicated replication forks moving in a leftward direction from the telomeres across the large replicon. This overall trend was apparent despite the residual activity of *ars609* in both Pu-seq strains and *ars606* in the ∆*6ori* Pu-seq strain (Figure 5.15, 5.15). A reduction in ribonucleotide signal across the large replicon was additionally observed in a ∆*4ori* Pu-seq strain (Figure 5.11). However, these data were noisier than those from the ∆*5ori* and ∆*6ori* Pu-seq strains, due to the lack of a complementary strain used for normalisation.

Of the two ways introduced to normalise data, the method used by Keszthelyi *et al*. was selected (Figure 5.10B, Method 2) (Keszthelyi *et al.* 2015).

This allowed normalisation on one strand of DNA between the mutant Polε and Polδ strains. An examination of the bias of mutant *pol2-M644G* and *pol3-L612G* to incorporate ribonucleotides indicated that they had the same preference order for substituting ribonucleotides into DNA (rC > rG > rA > rT, Figure 5.14). Therefore, by normalising on the same strand of DNA where the bases are the same, data were less subject to discrepancies due to homopolymer runs. Normalising to the other strand, as done by Clausen *et al*., may have led to bias in a poly(A) run, for example, as there is a higher propensity for rA to be incorporated than rT, regardless of polymerase (Figure 5.10B, Method 1) (Clausen *et al.* 2015). However, normalising to the other strand of DNA is more appropriate when DNA replication is performed by polymerases other than Polε and Polδ, because there are more than two polymerases contributing to replication. It was confirmed that the Pu-seq method of normalisation was appropriate by raw sequencing data which showed even coverage of reads in both mutant Polε and Polδ strains, with no reduction in reads incorporating ribonucleotides over the large replicon (Figure 5.13).

The ability of *ars606* to support plasmid replication after an extended incubation (Figure 5.18) was a new finding of this chapter. It had previously been reported that mutation of the ACS at an origin site to disrupt the consensus sequence led to complete inactivation of an origin (Hawkins *et al.* 2013b). Synchronous culture experiments (Figure 5.8), sort-seq data (Figure 4.22) and a plasmid assay with typical incubation time (2 days) (Figure 5.18) failed to detect any origin activity at *ars606*. However, the ability of polymerase-usage sequencing to measure the relative incorporation of ribonucleotides by Polε and Polδ showed replication initiation was still occurring at *ars606* at an efficiency of

approximately 50 % (measured as the height of the change in ribonucleotide signal). It was surprising that the plasmid activity of *ars606* was so weak when Pu-seq data showed an origin efficiency of 50 %. This may be due to the chromatin context in the genome being different to the DNA sequence alone in a plasmid context.

An origin efficiency of approximately 25 % was also observed at *ars609* (Figure 5.16). This was in contrast to the complete ablation of origin activity observed at *ars607*, *ars608* and *ars608.5* (Figure 5.12, 5.17). Examining the mutations introduced at each site (Table 4.1) does not highlight a difference in the location or extent of mutations introduced to the *ars606/ars609* and other origins. This chapter concludes that the impact of mutating the ACS at origins depends on the precise origin sequence and the mutations introduced. Included in this are considerations of surrounding sequence, such as potential secondary ORC binding sites. Chromatin context, for example nucleosome positioning may also contribute to the impact of mutating origin ACS sites. Future studies mutating different bases within the *ARS606* ACS are recommended to investigate a more effective origin inactivation that could be used in ongoing polymerase-usage studies on ChrVI.

Since residual origin activity was observed at both *ars606* and *ars609* it must be considered whether residual origin activity from *ars610* is responsible for the telomeric initiation observed in this chapter. Further, due to the core X element at the *ars610*, it was possible that recombination with other telomeres could have restored a wild-type core X origin. A PCR spanning the mutated X ARS and the *URA3* marker could be used to exclude this possibility. However, cells were frequently checked for growth on –URA plates, confirming the *URA3* marker had

not been lost. This discussion will consider explanations for replication initiation close to the telomeres, assuming that this is not due to residual origin activity.

Either a spatial or a sequence element located at the ends of chromosomes could facilitate replication. First a sequence element was considered. There are three forms of repetitive sequence found at telomeres; the X element, the Y element and telomere repeat sequences (TRSs), which have the general sequence $(C_{(1-3)}A)_n$. ChrVI right does not contain a Y element, whilst all telomeric sequences contain an X element. A circular plasmid containing the X element alone was shown to be unstable, although presence of the X element contributed to stability in the presence of short TRS repeats (Longtine *et al.* 1992). Yeast TRSs have been shown to stabilise replicative plasmids although they did not improve efficiency of replication (Longtine *et al.* 1992). Furthermore, Dershowitz *et al*. showed that in the absence of active origins the presence of TRS elements is not sufficient to facilitate replication of circular plasmids (Dershowitz *et al.* 2007). Therefore, a telomeric sequence element was not considered likely to lead to telomeric initiation.

In contrast, the topology of the chromosome has been shown to be essential for replication in the absence of active origins. Linearised versions of plasmids containing either only the inefficient late-firing origin *ARS308* or no active origins were shown to be replicated and maintained (Dershowitz *et al.* 2007). Therefore, it was hypothesised that the requirement for telomeric initiation was a linear chromosome end, not a sequence element contained within the telomere. In order to test this hypothesis, future work should focus on the creation of an internal large replicon alongside an artificial yeast telomere sequence, with internal $(C_{(1-3)}A)_n$ repeats. This would allow any initiation due to sequence to be

conserved, without the confounding variable of a chromosome end. One potential genomic locus is on ChrXIV, between *ARS1411* and *ARS1311.7*. Inactivation of confirmed origins *ARS1412*, *ARS1413*, *ARS1414* and *ARS1415* would result in an approximately 180 kb large replicon 170 kb away from the closest chromosome end (Nieduszynski *et al.* 2007; Siow *et al.* 2012).

Although the initiation site for telomeric replication was considered non-canonical, in that initiation would not involve an annotated active origin, the replication fork itself would not necessarily be abnormal. Previous studies have shown that replication initiation in the absence of active origins was ORC dependent (Bogenschutz *et al.* 2014; Dershowitz *et al.* 2007). Since ORC is the first protein recruited to ARS origins (Diffley *et al.* 1995), this suggested that a similar replication fork assembly pathway was occurring. Additionally, Pu-seq data showed that both Polε and Polδ were synthesising DNA at the telomeric initiation site, with Polε primarily incorporating ribonucleotides into the Crick (reverse) strand (Figure 5.12). It has been shown *in vitro* that at sites of DNA damage or replication fork restart both strands of DNA are initially synthesised by Polδ. An absence of Polδ bias over the large replicon indicated canonical replication fork progression moving leftwards from the telomere. A caveat to this conclusion is that Polδ leading strand synthesis is usually for short (less than one kb) regions, which are unlikely to be detected by Pu-seq (Yeeles *et al.* 2016).

Contrarily, telomeric initiation in both the ∆*4ori (m)* and ∆*6ori* synchronous culture experiments appeared to be limited (Figure 5.4, 5.8). At 70 % bulk genome replication, a clear increase of relative copy number was observed in the ∆*4ori (m)* and ∆*6ori* strains near the telomere. However, at 95 % bulk genome replication in the ∆*4ori (m)* strain this copy number had not further increased

closest to the telomere, and instead replication beyond a copy number of approximately 1.5 appeared dependent on replication from *ARS606* (Figure 5.4). In the ∆*6ori* strain, a small increase in relative copy number was observed close to the telomere between 70 % and 95 % bulk genome replication, but in the 95 % bulk genome replication sample this still only showed a copy number of approximately 1.5 (Figure 5.9). Similarly, in Pu-seq data both the ∆*5ori* and ∆*6ori* strains demonstrated a plateau in ribonucleotide signal at approximately 50 % synthesised by Polε on the Watson strand at the right end of ChrVI (Figure 5.12) This is despite a reduced ribonucleotide signal across the majority of the large replicon in the ∆*6ori* strain, relative to the ∆*5ori* strain. However, the ribonucleotide signal at the far right of ChrVI is confounded in both these strains by the residual activity of 'inactive' *ars609* (Figure 5.16) making meaningful comparisons more difficult.

Mathematical modelling of replication showed that replication initiating within the telomere and progressing leftwards fitted the observed ∆*5ori* Pu-seq strain data (Figure 5.19C). This model also resulted in a plateau of signal at around 50 % towards the right end of ChrVI, indicating that stochasticity in firing could lead to this effect (Figure 5.19D). However, it is important to note that the discrete initiation site hypothesised in this model had similar origin efficiency and mean replication time parameters to those used to simulate *ARS610*, located within the X element of the telomere. The maintenance of *URA3* in origin deletion strains was confirmed by selection on –URA plates. Furthermore, the absence of reads mapping to the reference genome at the end of ChrVI was confirmed in ∆*5ori* and ∆*6ori* Pu-seq sequencing reads. However, this study cannot rule out the possibility of canonical initiation from a site close to *ARS610*, along the lines

of the replication initiation observed by Bogenschutz *et al*. close to active origins after their deletion (Bogenschutz *et al.* 2014). This study also cannot exclude the possibility of a secondary ORC binding site close to *ars610* enabling residual origin activity, or the core X recombining with origin sequences in other core X elements to regain wild-type origin function.

In summary, this chapter presents synchronous culture and polymerase-usage evidence of a non-canonical form of replication initiation at the far right of ChrVI (telomeric initiation). It was concluded that telomeric initiation may be dependent on the linear topology of the ChrVI end and further studies using a discrete inter-chromosomal large replicon are recommended. The nature of replication forks initiating in the telomere initiation was considered. The literature suggests replication initiation in the presence of inactivated origins remains ORC-dependent. Repeating the *Δ5ori* and *Δ6ori* synchronous culture and Pu-seq experiments in an *orc2-1* temperature sensitive mutant to see whether telomeric initiation is still observed would be one way to investigate this further. However, the limited telomeric initiation suggests other mechanisms may be involved in this form of replication initiation.

6 Discussion and conclusions

6.1 Discussion

This thesis primary aim was to examine the regulation of DNA replication origins (Chapter 3) and the consequences of de-regulated DNA replication origins for DNA replication in the eukaryotes *S. cerevisiae* (Chapters 4 and 5). *S. cerevisiae* is an ideal model to study the regulation of origins due to the definition of origins at the level of DNA sequence (Nieduszynski *et al.* 2006), their well characterised nature (e.g. in the OriDB database) (Nieduszynski *et al.* 2007; Siow *et al.* 2012) and conservation of the replication initiation pathway between yeast and higher eukaryotes, such as humans (Bell & Dutta 2002). Until a few years ago, the replication initiation pathway, which begins with ORC binding to the origin ACS site, was considered to be a stepwise activation process (Yeeles *et al.* 2015). DDK then subsequent CDK phosphorylation of MCM helicase and firing factors respectively was thought to allow assembly of the active CMG helicase on DNA and origin firing (for review see Nishitani & Lygerou, 2002).

However, a wealth of recent literature outlined the importance of Rif1-PP1 in opposing the actions of DDK kinase (Davé *et al.* 2014; Hiraga *et al.* 2014; Mattarocci *et al.* 2014; Poh *et al.* 2014), making it clear that phosphatases also play an important role in controlling DNA replication origin firing. This thesis therefore investigated a putative role for another phosphatase subunit, Rts1, in regulation of origin firing (Chapter 3). *RTS1* had previously been reported to have a positive genetic interaction with the temperature-sensitive (ts-) ORC mutant *orc2*-1 (Costanzo *et al.* 2010). A panel of double mutant Δ*rts1* ts-replication factor yeast was assembled (section 3.2) to investigate genetic interactions between

RTS1 and replication factors. Deletion of *RTS1* in combination with mutations in origin licensing factor genes, *ORC* and *CDC6*, showed little or no genetic interaction (section 3.3, 3.4). This provided no genetic evidence for Rts1-PP2A controlling DNA replication origin licensing. Further, no role for Rts1 in opposing DDK phosphorylation was found (section 3.5). However, some level of increased temperature-sensitivity phenotype was observed when *RTS1* was deleted alongside ts-mutations in late-acting DNA replication firing factors Cdc45 and Dpb11, alluding to a potential negative interaction (section 3.6).

Additionally, an increased requirement for *RTS1* in a *cdc9-1* background was found (section 3.7). It can be speculated that a functional overlap between Rts1-PP2A and Cdc9 may exist via replication fork progression mechanisms. Rts1 may recruit PP2A during BIR, or the DDR, in response to DSBs. A caveat to the findings of sections 3.6 and 3.7 was the observation that deletion of a control gene, *TRP1*, also led to a small increase in the temperature-sensitivity of replication factor mutant strains (section 3.8). Nevertheless, the extent of ts-phenotype in Δ*trp1* ts-replication factor strains was not as pronounced as that observed in Δ*rts1 cdc9-1* and Δ*rts1 cdc45-27* strains.

Therefore, this study reported novel negative genetic interactions between *RTS1* and both the replication firing factor *CDC45*, and the *CDC9* ligase. A negative genetic interaction between *RTS1* and the replication factor *DPB11* was also observed, although this interaction was subsequently published in a large-scale genetic screen by Costanzo *et al*., after the completion of this chapter (Costanzo *et al.* 2016). Finally, a negative genetic interaction between *RTS1* and the replication factor *ORC2* was observed, suggesting the previous positive interaction reported by Costanzo *et al*. was a false positive – this was likely given

the previous observation was made in a high-throughput screen (Costanzo *et al.* 2010). It was concluded that any role played by Rts1 in DNA replication initiation occurred after the DDK step, and was facilitative. A potential role for Rts1 in promoting DNA elongation was also highlighted, which warrants further study. This work was published in Wellcome Open Research (Appendix 8.2) (Wallis & Nieduszynski 2018).

The second part of this thesis (Chapters 4 and 5) examined the consequences of de-regulating canonical DNA replication by inactivating adjacent origins to create a large replicon, greater than 100 kb. The origin *ARS608.5* on chromosome VI was confirmed in OriDB prior to this study but the precise ACS location was not known. Four putative ACS sites were identified by Amy Wolstenholme. Of these four, one was found to be essential for plasmid maintenance in yeast. Vamsi Gali (Donaldson laboratory) mutated the site using CRISPR/Cas9 (personal communication) and the mutated ACS site, *acs608.5*, was found to have reduced origin activity (section 4.3.2). Therefore, this thesis reported the novel finding of the precise functional origin sequence of *ARS608.5* (ChrVI: 253452-253466) and an inactivation with reduced origin activity.

It was found that the possession of an unusually large genomic replicon on ChrVI did not alter the mean generation time of cells (section 4.4). This was in agreement with previous studies which generated a large replicon on the left arm of ChrVI (Bogenschutz *et al.* 2014). In contrast, an unusually large replicon on ChrIII was found to slow cell growth (section 4.2.1). The potential for the ChrVI replicon to be replicated by replicating forks initiating at early-firing neighbouring origins was suggested to explain this difference in mean generation times of cells with large ChrIII or ChrVI replicons. Incomplete DNA replication during

metaphase leads to cell cycle arrest in a Rad9-dependent pathway (Ivanova *et al.* 2018). Therefore, a delay in DNA replication in the ∆*5ori* disome strain (relative to completion of replication in the ∆*0ori* disome strain) which prevented completion of DNA replication prior to metaphase could result in Rad9 activation and an extended cell cycle. Rad9 activation is CDK dependent, and leads to cell cycle delays in budding yeast with ongoing mitotic DNA replication (Ivanova *et al.* 2018; Wang *et al.* 2012). In contrast, in the ChrVI large replicon the mean generation time was not extended. This suggested DNA replication of the replicon was able to complete prior to metaphase and no Rad9 checkpoint was activated. It would be interesting to investigate whether a genetic interaction between ∆*6ori* and ∆*rad9* exists.

An important caveat to this conclusion is that this thesis examined only the overall mean generation times of cells. The S phase is defined as the phase during which DNA replication is occurring (Hartwell 1974). Since DNA replication was not complete across the ChrVI replicon in cells sorted by FACS into the 'G2' fraction, it could be argued that these were not true G2 cells. In order to understand the full impact of a large replicon on cell cycle progression in *S. cerevisiae* it would be important to examine the length of time yeast cells spent in each cell cycle phase. Inspection of flow cytometry profiles by FlowJo did not indicate a difference in the proportions of G1, S and G2 phase cells in the asynchronous cell populations of ∆*0ori,* ∆*4ori,* ∆*5ori* and ∆*6ori* yeast (data not shown). However, these data only encapsulate the bulk DNA content of cells, which would not have the sensitivity to detect ongoing replication at ChrVI right.

Cell fractions isolated by FACS (ChrVI) or cells from synchronous culture experiments (ChrIII) were analysed by ddPCR and/or deep sequencing to

investigate relative copy number of DNA across large replicons. Large replicons on both ChrVI (sections 4.6, 4.7) and ChrIII (section 4.2.2, Carolin Müller) were shown to delay local DNA replication. However, the delay to replication on ChrVI was less pronounced than the delay observed on ChrIII. This was also proposed to be due to neighbouring active origins (either *ARS606* or *ARS605*) on ChrVI which were capable of initiating replication forks to traverse the replicon early in S phase. In contrast, on the ChrIII disome the remaining origins were dormant, and late-firing (*ARS300-304*). However, both ChrVI and ChrIII replicons had not fully completed replication at the 'G2' stage of the cell cycle, where the majority of the genome was replicated. This suggested ongoing replication was occurring beyond the canonical S phase (although this is hard to define, as discussed above), and potentially concurrent with the start of mitosis.

During this study, a collaborator (Shin-ichoro Hiraga) investigated an independent strain with a large replicon on ChrVI, *Δ4ori (m)*. He found, using synchronous release experiments, that replication initiation could be observed at the far right on ChrVI after 70 % bulk DNA replication had occurred (section 5.2). This study showed that the late-firing initiation observed via synchronous release experiments could not be observed via sort-seq. Therefore, synchronous release experiments were used to analyse DNA replication of *Δ5ori* and *Δ6ori* strains later during the cell cycle than sort-seq facilitated (section 5.3). At 70 % bulk DNA replication both *Δ5ori* and *Δ6ori* strains also showed replication initiation at the far right on ChrVI, termed telomeric initiation.

Telomeric replication initiation was shown to be capable of progression in the *Δ6ori* strain. In the *Δ6ori* strain, at the time points of 70 % and 95 % bulk DNA

replication, telomeric initiation had a peak value of close to 25 % and 50 % replicated respectively. Contrastingly, the ∆*4ori (m)* strain indicated that the initiation of telomeric replication was limited, with the bulk of DNA replication across the replicon dependent on forks initiating at *ARS606*. In the ∆*4ori (m)* strain at the time points of both 70 % and 95 % bulk DNA replication, telomeric initiation had a peak value of close to 50 % replicated (1.5 relative copy number). Therefore, although telomeric initiation appears capable of processivity, this collective data did not show it progressing above 50 % replicated. Possible reasons for this include this study having not examined the correct time point to observe telomeric initiation at a more advanced stage. However, examination of Pu-seq data also indicated a mechanism for telomeric initiation where replication forks were initiating in the telomere in approximately 50 % of cells, as discussed below.

Regardless of the normalisation method used for ribonucleotide incorporation signal (discussed in detail in Chapter 5), it was shown in ∆*4ori*, ∆*5ori* and ∆*6ori* Pu-seq strains that the Watson (forward) strand of DNA was not primarily replicated by Polε across large replicons on ChrVI (sections 5.4, 5.5). Polε replication on the Crick strand and Polδ replication on the Watson strand in ∆*5ori* and ∆*6ori* Pu-seq strains indicated replication forks moving leftwards away from the telomere. As observed in the synchronous release experiments, it appeared that approximately 50 % of DNA at the far end of ChrVI was replicated by leftwards-moving forks.

A potential limitation of the Pu-seq data in ∆*5ori* and ∆*6ori* Pu-seq strains was that the levels of Polε and Polδ were considered only in respect to one another. However, the genome-wide coverage of reads and ribonucleotide

incorporation levels (section 5.5) indicated that there was no reduction in Polε or Polδ usage, or their propensity to incorporate ribonucleotides over the replicon on ChrVI. Therefore, it was concluded that the replicon on ChrVI was replicated primarily by replication forks containing both Polε and Polδ as opposed to Polδ/ Polδ replication forks resulting from restart after DNA damage (Yeeles *et al.* 2016), or alternative translesion polymerases being favoured (Goodman & Woodgate 2013).

A second limitation of the Pu-seq data collected was the incomplete inactivation of *ARS606* and *ARS609*. The published inactivation of *ARS606* (*ars606*) was shown to have activity in a plasmid assay after an extended incubation, and was observed in Pu-seq data to be capable of replication fork initiation at a reduced efficiency (section 5.6). Similarly, origin activity at *ars609* was observed, in the same genomic location as *ARS609*, although with reduced origin efficiency. A novel finding of this thesis was therefore that the previously described *ars606* and *ars609* variants do not fully prevent replication initiation, unlike the complete ablation of origin activity seen at *ars607*. It was concluded that it will be important to consider the context of the ACS and the precise base pairs mutated sider in future origin inactivation experiments. Origin context includes the potential for secondary ACS sites in the surrounding DNA, capable of binding ORC (Theis & Newlon 2001). In many such experiments, a delay to replication timing is the desired outcome and the *ars606* or *ars609* mutant origins are suitable for this use. However, in studies such as ORC-ChIP, where protein binding is examined, complete inactivation of the origin would be desired for meaningful analysis. Future studies should first check residual origin activity by a method such as Pu-seq before making conclusions based on an origin

inactivation confirmed by replication timing methods. Unfortunately, the residual activity of *ars609* masked the full extent of leftward-moving replication forks in Δ*5ori* and Δ*6ori* Pu-seq strains.

Nevertheless, both Pu-seq and synchronous release data indicated telomeric initiation occurred at the far right of ChrVI. Both Pu-seq and synchronous release data also supported a model where telomeric initiation was not responsible for more than approximately half of all replication at the far right of ChrVI. Mathematical modelling indicated that canonical replication forks initiating at the telomere mimicked this finding of the Pu-seq data, due the stochastic nature of replication fork firing (section 5.7). However, modelling relies on inputting values for the telomeric initiation site and is therefore subject to assumptions (for example, an assumption was made that telomeric initiation was late-firing). Consequently, this thesis considered biological mechanisms by which telomeric initiation could be occurring.

One form of replication initiation known to use only one strand of DNA is Break-Induced Replication (BIR) (Donnianni & Symington 2013). During BIR a single-stranded D-loop is formed which proceeds uni-directionally towards telomeres (Morrow *et al.* 1997). BIR has been shown to be capable of replicating DNA tracts greater than 100 kb, which is the length of the replicon introduced onto ChrVI in the Δ*5ori* strain (Donnianni & Symington 2013). However, BIR was more efficient the closer the site of initiation was to the end of the chromosome (Donnianni & Symington 2013). This form of replication could explain the increase in relative copy number seen in synchronous release experiments closer to the telomere without requiring initiation to occur solely in the telomere. Sites initiating

closer to the telomere are more efficient, and so would have greater copy number in this model.

Since BIR is also dependent on chromosome ends, this model was supported by the requirement for origin-less fragments to be linear (Dershowitz *et al.* 2007). Previous studies have examined the capacity of both telomere repeat sequences (TRSs) and the linear topology of telomeres to support plasmid replication (Dershowitz *et al.* 2007; Longtine *et al.* 1992). It was shown that in order to replicate in the absence of canonical active origins, DNA fragments required a linear topology. Contrastingly, the TRS repeat elements alone could not support plasmid replication (Longtine *et al.* 1992).

However, both Rad51-dependent and Rad51-independent BIR require Rad52 (Davis & Symington 2004). It was shown by Theis *et al*. that *RAD52* and other genes in its epistasis group were not required for the maintenance of origin-less ChrIII fragments in *S. cerevisiae* (Theis *et al.* 2010). This seemed to preclude the involvement of BIR in origin-less fragment maintenance. Further, BIR has been shown to be highly mutagenic (Smith *et al.* 2007) with 1000-fold higher mutation rates (Deem *et al.* 2011) due to the exposure of ssDNA on the lagging strand (Saini *et al.* 2013). Inspection of next-generation sequencing reads mapping across the replicon did not show a decrease in reads able to map, nor did visual inspection show high levels of mutagenesis (data not shown), although visual inspection was unlikely to have the sensitivity required to detect this effect.

Despite these contradictions, two proteins involved in BIR-like processes were found to be essential for origin-less fragment maintenance in *S. cerevisiae*: Sgs1 and Pol32 (Theis *et al.* 2010). Sgs1 is a homologue of the human BLM helicase (Yamagata *et al.* 1998), which is involved in resection of DSBs and

homologous recombination (HR) mediated repair (Gravel *et al.* 2008). Pol32 is a Polδ subunit, and a homologue of the human PolD3 subunit (Costantino *et al.* 2014). Normally non-essential, Pol32 is required for BIR (Lydeard *et al.* 2007). Additionally, it has been shown in human cells that the mutagenic effects of BIR can be suppressed by Mus81, which is capable of converting D-loops into full replication forks (Mayle *et al.* 2015). The latter study indicated that, where possible, once replication initiation by BIR had occurred the cell preferentially switched to replication via a full replication fork. Although Mus81 was not found to be essential for origin-less fragment maintenance, Mms4 was found to be essential (Theis *et al.* 2010). Mms4 can form a complex with Mus81, and Mms4-Mus81 has been found to function at stalled replication forks and have functional overlap with Sgs1 (Kaliraman *et al.* 2001). Therefore, it was considered that a BIR-like pathway, as opposed to canonical BIR which requires *RAD52*, may also be involved in telomeric initiation.

BIR-like pathways have been implicated in both Type I and Type II alternative lengthening of telomeres (ALT) pathways in telomerase-deficient cells (Dilley *et al.* 2016; Roumelioti *et al.* 2016). Both forms of ALT require Pol32, like telomeric initiation (Lydeard *et al.* 2007). Neither ALT (Chen *et al.* 2001), nor origin-less fragment maintenance is *Rad51*-dependent (Theis *et al.* 2010). However, both Type I and Type II replication were *RAD52* dependent (Chen *et al.* 2001), suggesting a form of ALT converting into a replication fork could not account for the telomeric initiation observed here. Further, a detailed genetic analysis of genetic requirements for ALT Type II in *S. cerevisiae* did not overlap with the genetic requirements for origin-less fragment maintenance, beyond the *SGS1* and *POL32* genes already discussed (Figure 6.1) (Theis *et al.* 2010).

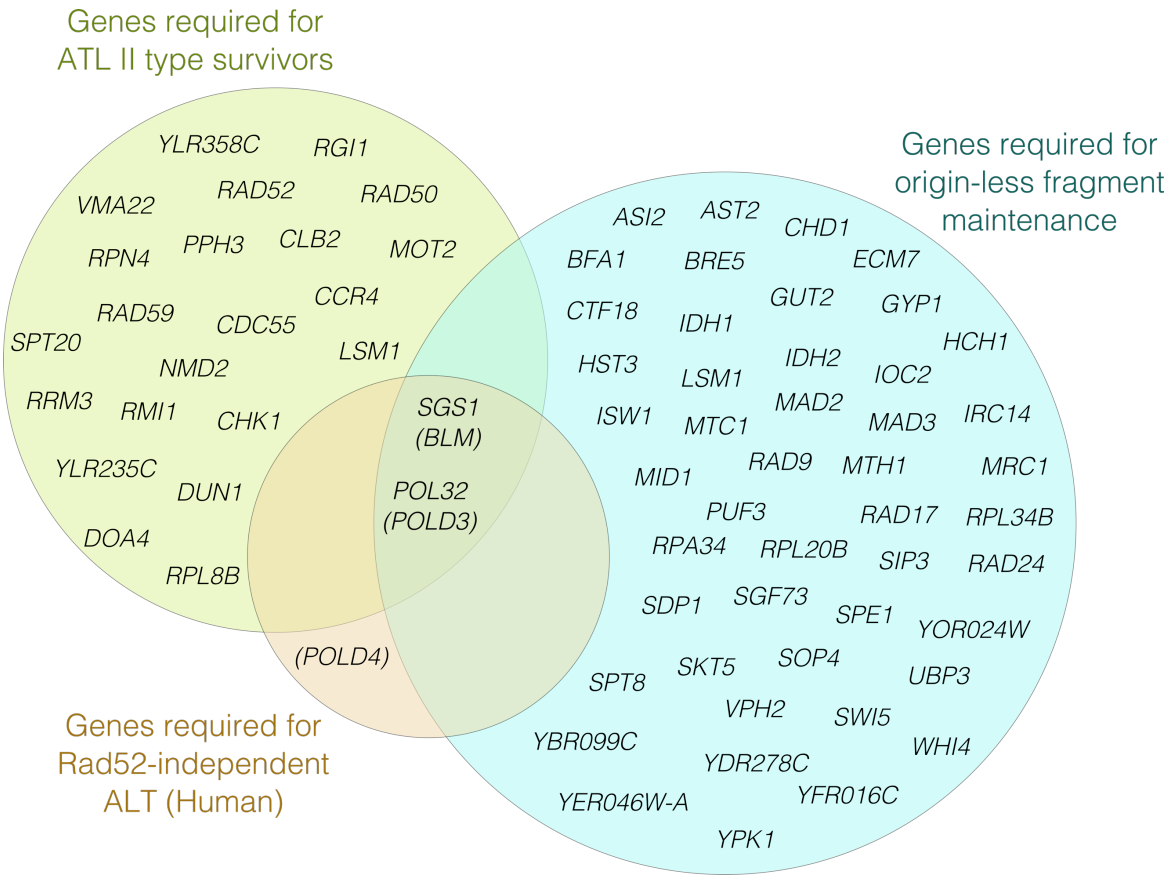


Figure 6.1 The genetic requirements of Rad52-independent ALT and origin-less fragment maintenance overlap with those of ALT II. *Genes identified in S. cerevisiae as essential for Alternative-Lengthening of Telomeres (ALT) II type survivors (Rad51-independent) (Green)* (Lydeard *et al.* 2007; van Mourik *et al.* 2018) *or essential for origin-less fragment maintenance (Blue)* (Theis *et al.* 2010) *are highlighted. Genes identified in human U20S cells as being involved in a recently discovered Rad51- and Rad52-independent ALT process (Orange, brackets) and, if applicable, their yeast homologs are shown (Orange)* (Zhang *et al.* 2019).

Intriguingly, this year an alternative BIR-like ALT pathway was discovered in human cells which does not require *RAD51* or *RAD52* (Zhang *et al.* 2019). This pathway was dependent on BLM and PolD3/PolD4 (Sgs1 and Pol32 in yeast) meaning the only currently known genetic requirements are those identified as overlapping between ALT and origin-less fragment maintenance (Figure 6.1) (Zhang *et al.* 2019). Further work is needed to elucidate the mechanism involved in *RAD52*-independent BIR, however, it is tantalising to speculate that this or a similar process could be occurring at telomere ends in cells with a large replicon, such as the ChrVI telomeric initiation described in this thesis.

It is important to note that this thesis cannot rule out the possibility that there is residual activity of *ARS610*, due to incomplete inactivation, secondary ACS sites or the core X recombining with other telomeric core X elements and regaining origin function. In order to test this, a plasmid assay using DNA amplified from *ars610* could be carried out with an extended incubation time, similar to the plasmid assay used to determine *ars606* still had weak origin activity (section 5.6). Additionally, replication from non-canonical initiation sites which are below the detection threshold for population based methods has been found recently in *S. cerevisiae* (Müller *et al.* 2019). Although modelling indicated that random initiation could not explain the Pu-seq data (section 5.7) it was considered that non-site specific initiation could occur across the ChrVI right large replicon non-randomly. For example, transcription may interact with potential initiation sites by removing pre-RC proteins. Since less transcription occurs close to telomeres, this would result in higher initiation rates closer to the telomere. However, mathematical modelling of this scenario gave a similar result to that

described in section 5.7; any model with initiation occurring across the large replicon resulted in 100 % of forks moving rightwards at the end of ChrVI (271 kb).

6.2 Conclusions

This thesis aimed to investigate the regulation of DNA replication origins and consequences for cell growth and DNA replication when origins were de-regulated. It was shown that Rts1 did not antagonise origin firing. Contrastingly, this thesis found that *RTS1* had negative genetic interactions with *CDC9*, *DBP11* and *CDC45*. Secondly, two large replicons were studied, one in a yeast strain with a second ChrIII fragment and the other on the right arm of ChrVI. It was shown that, in the absence of active origins, DNA replication was delayed. However, this did not necessarily lead to a delayed cell cycle.

Thirdly, it was shown that increased relative copy number, indicating replication initiation, occurred during late-S phase when active origins were not present on the right arm of ChrVI. Polymerase-usage data indicated that canonical replication forks were able to progress leftwards into the large replicon. Some of the genetic requirements in the ChrIII *Δ5ori* fragment strain for origin-less fragment maintenance were similar to those required for BIR (Gravel *et al.* 2008; Lydeard *et al.* 2007). BIR can be non-mutagenic and converted to full replication forks by Mus81 (Mayle *et al.* 2015), and it was recently shown that a *RAD52*-independent form of BIR can be used for alternative lengthening of telomeres in human cells (Zhang *et al.* 2019).

Future studies to examine replication in the absence of active origins should include a non-telomeric large replicon, for example on ChrXIV, between *ARS1411* and *ARS1311.7*, containing an artificial yeast telomere sequence, with internal $(C_{(1-3)}A)_n$ repeats. This would enable separation of the sequence element of telomeres from the topology of a chromosome end. If BIR-like mechanisms were responsible for the observed telomeric initiation, it would be predicted that no non-canonical initiation would be observed from the artificial yeast telomere. Secondly, examination of single molecules would be ideal to determine how replication forks traverse the large replicon in more detail. D-NAscent, developed by the Nieduszynski laboratory, allows the quantification of BrdU incorporation into ultra-long single molecules to determine initiation sites on individual molecules (Müller *et al.* 2019).

In conclusion, this thesis has shown that when multiple adjacent DNA replication origins were inactivated an alternative form of replication initiation was able to increase relative copy number near the telomeres. The work presented here may indicate that a newly reported *RAD52*-independent ALT pathway in humans is also able to function in yeast. Regardless, it has been reported that in the absence of DNA replication origins a replicon on the right arm of ChrVI in *S. cerevisiae* was replicated by leftward moving replication forks which appeared to initiate at the far right of ChrVI, a phenomenon here termed telomeric initiation. It remains to be seen whether telomeric initiation occurs at all telomeres in the *S. cerevisiae* genome. Given that all telomeres replicate late in S phase it is possible that an alternative replication initiation pathway exists in order to ensure they are replicated prior to cell division. The importance of replicating telomeres was highlighted by a modelling study which predicted (then subsequently observed)

an increased occurrence of origins proximal to telomeres in *S. cerevisiae* (Newman *et al.* 2013).

7 Bibliography

Aggarwal, B. D., & Calvi, B. R. (2004). Chromatin regulates origin activity in Drosophila follicle cells. *Nature*, **430**, 372–376.

Agier, N., Delmas, S., Zhang, Q., Fleiss, A., Jaszczyszyn, Y., Van Dijk, E., Thermes, C., Weigt, M., Cosentino-Lagomarsino, M., & Fischer, G. (2018). The evolution of the temporal program of genome replication. *Nature Communications*, **9**, 2199.

Aguilera, A., & Garcia-Muse, T. (2012). R Loops: From transcription byproducts to threats to genome stability. *Molecular Cell*, **46**(2), 115–124.

Alabert, C., & Groth, A. (2012). Chromatin replication and epigenome maintenance. *Nature Reviews Molecular Cell Biology*, **13**, 153–167.

Alcasabas, A. A., Osborn, A. J., Bachant, J., Hu, F., Werler, P. J. H., Bousset, K., Furuya, K., Diffley, J. F. X., Carr, A. M., & Elledge, S. J. (2001). Mrc1 transduces signals of DNA replication stress to activate Rad53. *Nature Cell Biology*, **3**(11), 958–965.

Alexander, J. L., & Orr-Weaver, T. L. (2016). Replication fork instability and the consequences of fork collisions from rereplication. *Genes and Development*, **30**(20), 2241–2252.

Alver, R. C., Chadha, G. S., Gillespie, P. J., & Blow, J. J. (2017). Reversal of DDK-mediated MCM phosphorylation by Rif1-PP1 regulates replication initiation and replisome stability independently of ATR/Chk1. *Cell Reports*, **18**(10), 2508–2520.

Alvino, G. M., Collingwood, D., Murphy, J. M., Delrow, J., Brewer, B. J., & Raghuraman, M. K. (2007). Replication in hydroxyurea: It's a matter of time. *Molecular and Cellular Biology*, **27**(18), 6396–6406.

Aparicio, J. G., Viggiani, C. J., Gibson, D. G., & Aparicio, O. M. (2004). The Rpd3-Sin3 histone deacetylase regulates replication timing and enables intra-S origin control in Saccharomyces cerevisiae. *Molecular and Cellular Biology*, **24**(11), 4769–4780.

Aparicio, O. M., Stout, A. M., & Bell, S. P. (1999). Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication. *Proceedings of the*

National Academy of Sciences of the United States of America, **96**(16), 9130–9135.

Araki, H. (2016). Elucidating the DDK-dependent step in replication initiation. *The EMBO Journal*, **35**(9), 907–908.

Archambault, V., Ikui, A. E., Drapkin, B. J., & Cross, F. R. (2005). Disruption of mechanisms that prevent rereplication triggers a DNA damage response. *Molecular and Cellular Biology*, **25**(15), 6707–6721.

Arora, R., Lee, Y., Wischnewski, H., Brun, C. M., Schwarz, T., & Azzalin, C. M. (2014). RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat Commun*, **5**, 5220.

Asai, T., Imai, M., & Kogoma, T. (1994). DNA Damage-inducible replication of the Escherichia coli chromosome is initiated at separable sites within the minimal oriC. *Journal of Molecular Biology*, **235**, 1459–1469.

Asai, T., Sommer, S., Bailone, A., & Kogoma, T. (1993). Homologous recombination-dependent initiation of DNA replication from DNA damage-inducible origins in Escherichia coli. *The EMBO Journal*, **12**(8), 3287–3295.

Auesukaree, C., Damnernsawad, A., Kruatrachue, M., Pokethitiyook, P., Boonchird, C., Kaneko, Y., & Harashima, S. (2009). Genome-wide identification of genes involved in tolerance to various environmental stresses in Saccharomyces cerevisiae. *Journal of Applied Genetics*, **50**(3), 301–310.

Ausiannikava, D., & Allers, T. (2017). Diversity of DNA replication in the archaea. *Genes*, **8**(2), 56.

Aves, S. J., Liu, Y., & Richards, T. A. (2012). Evolutionary diversification of eukaryotic DNA replication machinery. *Subcellular Biochemistry*, **62**, 19–35.

Ayyagari, R., Gomes, X. V., Gordenin, D. A., & Burgers, P. M. J. (2003). Okazaki fragment maturation in yeast: I. Distribution of functions between FEN1 and DNA2. *Journal of Biological Chemistry*, **278**(3), 1618–1625.

Bairwa, N. K., Zzaman, S., Mohanty, B. K., & Bastia, D. (2010). Replication fork arrest and rDNA silencing are two independent and separable functions of the replication

terminator protein Fob1 of Saccharomyces cerevisiae. *Journal of Biological Chemistry*, **285**(17), 12612–12619.

Baker Brachmann, C., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., & Boeke, J. D. (1998). Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, **14**(2), 115–132.

Balakrishnan, L., & Bambara, R. A. (2013). Flap endonuclease 1. *Annual Review of Biochemistry*, **82**, 119–138.

Balestrini, A., Cosentino, C., Errico, A., Garner, E., & Costanzo, V. (2010). GEMC1 is a TopBP1-interacting protein required for chromosomal DNA replication. *Nature Cell Biology*, **12**(5), 484–491.

Bárcena, M., Ruiz, T., Donate, L. E., Brown, S. E., Dixon, N. E., Radermacher, M., & Carazo, J. M. (2001). The DnaB·DnaC complex: A structure based on dimers assembled around an occluded channel. *EMBO Journal*, **20**(6), 1462–1468.

Bastia, D., & Zaman, S. (2014). Mechanism and physiological significance of programmed replication termination. *Seminars in Cell and Developmental Biology*, **30**, 165–173.

Batrakou, D. G., Heron, E. D., & Nieduszynski, C. A. (2018). Rapid high-resolution measurement of DNA replication timing by droplet digital PCR. *Nucleic Acids Research*, **46**(19), e112.

Baxter, J., & Diffley, J. F. X. (2008). Topoisomerase II inactivation prevents the completion of DNA replication in budding yeast. *Molecular Cell*, **30**(6), 790–802.

Baxter, J., Sen, N., López Martínez, V., Monturus De Carandini, M. E., Schvartzman, J. B., Diffley, J. F. X., & Aragón, L. (2011). Positive supercoiling of mitotic DNA drives decatenation by topoisomerase II in eukaryotes. *Science*, **331**(6022), 1328–1332.

Bell, S. D. (2012). Archaeal Orc1/Cdc6 proteins. *Subcellular Biochemistry*, **62**, 59–69.

Bell, S. P., & Dutta, A. (2002). DNA replication in eukaryotic cells. *Annual Review of Biochemistry*, **71**, 333–374.

Bell, S. P., & Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature*, **357**(6374), 128–134.

Berbenetz, N. M., Nislow, C., & Brown, G. W. (2010). Diversity of eukaryotic DNA replication origins revealed by Genome-wide analysis of chromatin structure. *PLoS Genetics*, **6**(9), e1001092.

Berezney, R., Dubey, D. D., & Huberman, J. A. (2000). Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. *Chromosoma*, **108**(8), 471–484.

Bergman, L. W. (2001). Growth and maintenance of yeast. *Methods in Molecular Biology*, **177**, 9–14.

Bermejo, R., Doksani, Y., Capra, T., Katou, Y. M., Tanaka, H., Shirahige, K., & Foiani, M. (2007). Top1- and Top2-mediated topological transitions at replication forks ensure fork progression and stability and prevent DNA damage checkpoint activation. *Genes and Development*, **21**(15), 1921–1936.

Bermudez, V. P., Farina, A., Tappin, I., & Hurwitz, J. (2010). Influence of the human cohesion establishment factor Ctf4/AND-1 on DNA replication. *Journal of Biological Chemistry*, **285**(13), 9493–9505.

Bhowmick, R., Minocherhomji, S., & Hickson, I. D. (2016). RAD52 facilitates mitotic DNA synthesis following replication stress. *Molecular Cell*, **64**(6), 1117–1126.

Bi, X. (2014). Heterochromatin structure: Lessons from the budding yeast. *IUBMB Life*, **66**(10), 657–666.

Bickel, S. E., & Orr-Weaver, T. L. (1996). Holding chromatids together to ensure they go their separate ways. *BioEssays*, **18**(4), 293–300.

Bicknell, L. S., Walker, S., Klingseisen, A., Stiff, T., Leitch, A., Kerzendorfer, C., Martin, C. A., Yeyati, P., Al Sanna, N., Bober, M., Johnson, D., Wise, C., Jackson, A. P., O'Driscoll, M., & Jeggo, P. A. (2011). Mutations in ORC1, encoding the largest subunit of the origin recognition complex, cause microcephalic primordial dwarfism resembling Meier-Gorlin syndrome. *Nature Genetics*, **43**(4), 350–355.

Blake, D., Luke, B., Kanellis, P., Jorgensen, P., Goh, T., Penfold, S., Breitkreutz, B. J.,

Durocher, D., Peter, M., & Tyers, M. (2006). The F-box protein Dia2 overcomes replication impedance to promote genome stability in Saccharomyces cerevisiae. *Genetics*, **174**(4), 1709–1727.

Blow, J. J., & Dutta, A. (2005). Preventing re-replication of chromosomal DNA. *Nature Reviews Molecular Cell Biology*, **6**(6), 476–486.

Blow, J. J., & Ge, X. Q. (2008). Replication forks, chromatin loops and dormant replication origins. *Genome Biology*, **9**(12), 244.

Blow, J. J., Ge, X. Q., & Jackson, D. A. (2011). How dormant origins promote complete genome replication. *Trends in Biochemical Sciences*, **36**(8), 405–414.

Blow, J. J., & Gillespie, P. J. (2008). Replication licensing and cancer - A fatal entanglement? *Nature Reviews Cancer*, **8**(10), 799–806.

Blow, J. J., & Hodgson, B. (2002). Replication licensing - Defining the proliferative state? *Trends in Cell Biology*, **12**(2), 72–78.

Blow, J. J., & Tanaka, T. U. (2005). The chromosome cycle: Coordinating replication and segregation. Second in the cycles review series. *EMBO Reports*, **6**(11), 1028–1034.

Bochman, M. L., & Schwacha, A. (2008). The Mcm2-7 complex has in vitro helicase activity. *Molecular Cell*, **31**(2), 287–293.

Boemo, M. A., & Byrne, H. M. (2019). Mathematical modelling of a hypoxia-regulated oncolytic virus delivered by tumour-associated macrophages. *Journal of Theoretical Biology*, **461**, 102–116.

Boemo, M. A., Cardelli, L., & Nieduszynski, C. A. (2019). The Beacon Calculus: A formal method for the flexible and concise modelling of biological systems. *BioRxiv*, **579029**. doi:10.1101/579029

Bogenschutz, N. L., Rodriguez, J., & Tsukiyama, T. (2014). Initiation of DNA replication from non-canonical sites on an origin-depleted chromosome. *PLoS ONE*, **9**(12), 1–11.

Boos, D., Sanchez-Pulido, L., Rappas, M., Pearl, L. H., Oliver, A. W., Ponting, C. P., & Diffley, J. F. X. (2011). Regulation of DNA replication through Sld3-Dpb11

interaction is conserved from yeast to humans. *Current Biology*, **21**(13), 1152–1157.

Botstein, D., Chervitz, S. A., & Cherry, J. M. (1997). Yeast as a model organism. *Science*, **277**(5330), 1259–1260.

Bouton, A. H., & Smith, M. M. (1986). Fine-structure analysis of the DNA sequence requirements for autonomous replication of Saccharomyces cerevisiae plasmids. *Molecular and Cellular Biology*, **6**(7), 2354–2363.

Bowers, J. L., Randell, J. C. W., Chen, S., & Bell, S. P. (2004). ATP hydrolysis by ORC catalyzes reiterative Mcm2-7 assembly at a defined origin of replication. *Molecular Cell*, **16**(6), 967–978.

Bowman, G. D., & McKnight, J. N. (2017). Sequence-specific targeting of chromatin remodelers organizes precisely positioned nucleosomes throughout the genome. *BioEssays*, **39**(1), 1–8.

Brassinga, A. K. C., Siam, R., McSween, W., Winkler, H., Wood, D., & Marczynski, G. T. (2002). Conserved response regulator CtrA and IHF binding sites in the α-proteobacteria Caulobacter crescentus and Rickettsia prowazekii chromosomal replication origins. *Journal of Bacteriology*, **184**(20), 5789–5799.

Breier, A. M., Chatterji, S., & Cozzarelli, N. R. (2004). Prediction of Saccharomyces cerevisiae replication origins. *Genome Biology*, **5**(4), R22.

Brewer, B. J., & Fangman, W. L. (1987). The localization of replication origins on ARS plasmids in S. cerevisiae. *Cell*, **51**(3), 463–471.

Brewer, B. J., & Fangman, W. L. (1993). Initiation at closely spaced replication origins in a yeast chromosome. *Science*, **262**(5140), 1728–1731.

Brocard-Masson, C., & Dumas, B. (2006). The fascinating world of steroids: S. cerevisiae as a model organism for the study of hydrocortisone biosynthesis. *Biotechnology and Genetic Engineering Reviews*, **22**, 213–252.

Bruck, I., & Kaplan, D. L. (2011a). GINS and Sld3 compete with one another for Mcm2-7 and Cdc45 binding. *Journal of Biological Chemistry*, **286**(16), 14157–14167.

Bruck, I., & Kaplan, D. L. (2011b). Origin single-stranded DNA releases Sld3 protein from

the Mcm2-7 complex, allowing the GINS tetramer to bind the Mcm2-7 complex. *Journal of Biological Chemistry*, **292**(24), 10318.

Bruck, I., & Kaplan, D. L. (2015). The Dbf4-Cdc7 kinase promotes Mcm2-7 ring opening to allow for single-stranded DNA extrusion and helicase assembly. *Journal of Biological Chemistry*, **290**(2), 1210–1221.

Bullock, P. A., Seo, Y. S., & Hurwitz, J. (1991). Initiation of simian virus 40 DNA synthesis in vitro. *Molecular and Cellular Biology*, **11**(5), 2350–2361.

Buongiorno-Nardelli, M., Micheli, G., Carrĩ, M. T., & Marilley, M. (1982). A relationship between replicon size and supercoiled loop domains in the eukaryotic genome. *Nature*, **298**(5869), 100–102.

Burgers, P. M. J. (2009). Polymerase dynamics at the eukaryotic DNA replication fork. *Journal of Biological Chemistry*, **284**(7), 4041–4045.

Burgers, P. M. J., Koonin, E. V., Bruford, E., Blanco, L., Burtis, K. C., Christman, M. F., Copeland, W. C., Friedberg, E. C., Hanaoka, F., Hinkle, D. C., Lawrence, C. W., Nakanishi, M., Ohmori, H., Prakash, L., Prakash, S., Reynaud, C. A., Sugino, A., Todo, T., Wang, Z., Weill, J. C., & Woodgate, R. (2001). Eukaryotic DNA polymerases: Proposal for a revised nomenclature. *Journal of Biological Chemistry*, **276**(47), 43487–43490.

Burgers, P. M. J., & Kunkel, T. A. (2017). Eukaryotic DNA replication fork. *Annual Review of Biochemistry*, **86**, 417–438.

Burland, V., Plunkett, G., Daniels, D. L., & Blattner, F. R. (1993). DNA sequence and analysis of 136 kilobases of the Escherichia coli genome: organizational symmetry around the origin of replication. *Genomics*, **16**(3), 551–561.

Caddle, M. S., & Calos, M. P. (1994). Specific initiation at an origin of replication from Schizosaccharomyces pombe. *Molecular and Cellular Biology*, **14**(3), 1796–1805.

Cadoret, J. C., Meisch, F., Hassan-Zadeh, V., Luyten, I., Guillet, C., Duret, L., Quesneville, H., & Prioleau, M. N. (2008). Genome-wide studies highlight indirect links between human replication origins and gene regulation. *Proceedings of the*

National Academy of Sciences of the United States of America, **105**(41), 15837–15842.

Calzada, A., Hodgson, B., Kanemaki, M., Bueno, A., & Labib, K. (2005). Molecular anatomy and regulation of a stable replisome at a paused eukaryotic DNA replication fork. *Genes and Development*, **19**(16), 1905–1919.

Castermans, D., Somers, I., Kriel, J., Louwet, W., Wera, S., Versele, M., Janssens, V., & Thevelein, J. M. (2012). Glucose-induced posttranslational activation of protein phosphatases PP2A and PP1 in yeast. *Cell Research*, **22**(6), 1058–1077.

Cayrou, C., Coulombe, P., Puy, A., Rialle, S., Kaplan, N., Segal, E., & Méchali, M. (2012). New insights into replication origin characteristics in metazoans. *Cell Cycle*, **11**(4), 658–667.

Cayrou, C., Coulombe, P., Vigneron, A., Stanojcic, S., Ganier, O., Peiffer, I., Rivals, E., Puy, A., Laurent-Chabalier, S., Desprat, R., & Méchali, M. (2011). Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features. *Genome Research*, **21**(9), 1438–1449.

Cerritelli, S. M., Frolova, E. G., Feng, C., Grinberg, A., Love, P. E., & Crouch, R. J. (2003). Failure to produce mitochondrial DNA results in embryonic lethality in Rnaseh1 null mice. *Molecular Cell*, **11**(3), 807–815.

Chan, Y. A., Aristizabal, M. J., Lu, P. Y. T., Luo, Z., Hamza, A., Kobor, M. S., Stirling, P. C., & Hieter, P. (2014). Genome-wide profiling of yeast DNA:RNA hybrid prone sites with DRIP-Chip. *PLoS Genetics*, **10**(4), e1004288.

Chastain, P. D., Makhov, A. M., Nossal, N. G., & Griffith, J. (2003). Architecture of the replication complex and DNA loops at the fork generated by the bacteriophage T4 proteins. *Journal of Biological Chemistry*, **278**(23), 21276–21285.

Chen, C. L., Duquenne, L., Audit, B., Guilbaud, G., Rappailles, A., Baker, A., Huvet, M., D'Aubenton-Carafa, Y., Hyrien, O., Arneodo, A., & Thermes, C. (2011). Replication-associated mutational asymmetry in the human genome. *Molecular Biology and Evolution*, **28**(8), 2327–2337.

Chen, Q., Ijpma, A., & Greider, C. W. (2001). Two survivor pathways that allow growth in the absence of telomerase are generated by distinct telomere recombination events. *Molecular and Cellular Biology*, **21**(5), 1819–1827.

Chen, S., & Bell, S. P. (2011). CDK prevents Mcm2-7 helicase loading by inhibiting Cdt1 interaction with Orc6. *Genes and Development*, **25**(4), 363–372.

Chen, S., De Vries, M. A., & Bell, S. P. (2007). Orc6 is required for dynamic recruitment of Cdt1 during repeated Mcm2-7 loading. *Genes and Development*, **21**(22), 2897–2907.

Cherry, J. M., Adler, C., Ball, C., Chervitz, S. A., Dwight, S. S., Hester, E. T., Jia, Y., Juvik, G., Roe, T., Schroeder, M., Weng, S., & Botstein, D. (1998). SGD: Saccharomyces genome database. *Nucleic Acids Research*, **26**(1), 73–79.

Chilkova, O., Stenlund, P., Isoz, I., Stith, C. M., Grabowski, P., Lundström, E. B., Burgers, P. M., & Johansson, E. (2007). The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA. *Nucleic Acids Research*, **35**(19), 6588–6597.

Chou, D. M., Petersen, P., Walter, J. C., & Walter, G. (2002). Protein phosphatase 2A regulates binding of Cdc45 to the prereplication complex. *Journal of Biological Chemistry*, **277**(43), 40520–40527.

Chow, T. T., Zhao, Y., Mak, S. S., Shay, J. W., & Wright, W. E. (2012). Early and late steps in telomere overhang processing in normal human cells: The position of the final RNA primer drives telomere shortening. *Genes and Development*, **26**(11), 1167–1178.

Chowdhury, A., Liu, G., Kemp, M., Chen, X., Katrangi, N., Myers, S., Ghosh, M., Yao, J., Gao, Y., Bubulya, P., & Leffak, M. (2010). The DNA Unwinding Element Binding Protein DUE-B Interacts with Cdc45 in Preinitiation Complex Formation. *Molecular and Cellular Biology*, **30**(6), 1495–1507.

Choy, J. S., Mishra, P. K., Au, W. C., & Basrai, M. A. (2012). Insights into assembly and

regulation of centromeric chromatin in Saccharomyces cerevisiae. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, **1819**(7), 776–783.

Clausen, A. R., Lujan, S. A., Burkholder, A. B., Orebaugh, C. D., Williams, J. S., Clausen, M. F., Malc, E. P., Mieczkowski, P. A., Fargo, D. C., Smith, D. J., & Kunkel, T. A. (2015). Tracking replication enzymology in vivo by genome-wide mapping of ribonucleotide incorporation. *Nature Structural and Molecular Biology*, **22**(3), 185–191.

Clausen, A. R., Zhang, S., Burgers, P. M., Lee, M. Y., & Kunkel, T. A. (2013). Ribonucleotide incorporation, proofreading and bypass by human DNA polymerase δ. *DNA Repair*, **12**(2), 121–127.

Clyne, R. K., & Kelly, T. J. (1995). Genetic analysis of an ARS element from the fission yeast Schizosaccharomyces pombe. *The EMBO Journal*, **14**(24), 6348–6357.

Cohen-Fix, O., & Koshland, D. (1997). The metaphase-to-anaphase transition: Avoiding a mid-life crisis. *Current Opinion in Cell Biology*, **9**(6), 800–806.

Coker, J. A., DasSarma, P., Capes, M., Wallace, T., McGarrity, K., Gessler, R., Liu, J., Xiang, H., Tatusov, R., Berquist, B. R., & DasSarma, S. (2009). Multiple replication origins of Halobacterium sp. strain NRC-1: Properties of the conserved orc7-dependent oriC1. *Journal of Bacteriology*, **191**(16), 5253–5261.

Collins, F. S., Lander, E. S., Rogers, J., & Waterson, R. H. (2004). Finishing the euchromatic sequence of the human genome. *Nature*, **431**(7011), 931–945.

Collins, S. R., Miller, K. M., Maas, N. L., Roguev, A., Fillingham, J., Chu, C. S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M., Ding, H., Xu, H., Han, J., Ingvarsdottir, K., Cheng, B., Andrews, B., Boone, C., Berger, S. L., Hieter, P., Zhang, Z., Brown, G. W., Ingles, C. J., Emili, A., Allis, C. D., Toczyski, D. P., Weissman, J. S., Greenblatt, J. F., & Krogan, N. J. (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature*, **446**(7137), 806–10.

Cosgrove, A. J., Nieduszynski, C. A., & Donaldson, A. D. (2002). Ku complex controls

the replication time of DNA in telomere regions. *Genes and Development*, **16**(19), 2485–2490.

Costa, A., Ilves, I., Tamberg, N., Petojevic, T., Nogales, E., Botchan, M. R., & Berger, J. M. (2011). The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nature Structural and Molecular Biology*, **18**(4), 471–477.

Costantino, L., Sotiriou, S. K., Rantala, J. K., Magin, S., Mladenov, E., Helleday, T., Haber, J. E., Iliakis, G., Kallioniemi, O. P., & Halazonetis, T. D. (2014). Break-induced replication repair of damaged forks induces genomic duplications in human cells. *Science*, **343**(6166), 88–91.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E. D., Sevier, C. S., Ding, H., Koh, J. L. Y., Toufighi, K., Mostafavi, S., Prinz, J., St Onge, R. P., VanderSluis, B., Makhnevych, T., Vizeacoumar, F. J., Alizadeh, S., Bahr, S., Brost, R. L., Chen, Y., Cokol, M., Deshpande, R., Li, Z., Lin, Z.-Y., Liang, W., Marback, M., Paw, J., San Luis, B.-J., Shuteriqi, E., Tong, A. H. Y., van Dyk, N., Wallace, I. M., Whitney, J. A., Weirauch, M. T., Zhong, G., Zhu, H., Houry, W. A., Brudno, M., Ragibizadeh, S., Papp, B., Pál, C., Roth, F. P., Giaever, G., Nislow, C., Troyanskaya, O. G., Bussey, H., Bader, G. D., Gingras, A.-C., Morris, Q. D., Kim, P. M., Kaiser, C. A., Myers, C. L., Andrews, B. J., Boone, C., Dixon, S. J., Costanzo, M., Baryshnikova, A., Andrews, B., Boone, C., Hartwell, L., Tong, A. H., Tong, A. H., Mani, R., Onge, R. P. St., Hartman, J. L., Giaever, G., Roth, F. P., Onge, R. P. S., Segrè, D., Deluna, A., Church, G. M., Kishony, R., Huber, A., Chen, E. J., Kaiser, C. A., Jonikas, M. C., Metzger, M. B., Michaelis, S., Leidel, S., Rahl, P. B., Chen, C. Z., Collins, R. N., Esberg, A., Huang, B., Johansson, M. J., Byström, A. S., Naumanen, T., Johansen, L. D., Coffey, E. T., Kallunki, T., Johansen, L. D., Levy, S. F., Siegal, M. L., Levchenko, A., Kim, P. M., Lu, L. J., Xia, Y., Gerstein, M. B., Fraser, H. B., Wall, D. P., Hirsh, A. E., Pál, C., Papp, B., Hurst, L. D., Kim, P. M., Sboner, A., Xia, Y., Gerstein, M., Lehner, B., Crombie, C., Tischler, J., Fortunato, A., Fraser, A. G., Hillenmeyer, M. E., Dunham, M. J., Gavin, A. C., Krogan, N. J., Yu, H., Tarassov,

K., Visser, J. A. de, Meiklejohn, C. D., Hartl, D. L., Parsons, A. B., & Sevier, C. S. (2010). The genetic landscape of a cell. *Science*, **327**(5964), 425–431.

Costanzo, M., VanderSluis, B., Koch, E. N., Baryshnikova, A., Pons, C., Tan, G., Wang, W., Usaj, M., Hanchard, J., Lee, S. D., Pelechano, V., Styles, E. B., Billmann, M., van Leeuwen, J., van Dyk, N., Lin, Z.-Y., Kuzmin, E., Nelson, J., Piotrowski, J. S., Srikumar, T., Bahr, S., Chen, Y., Deshpande, R., Kurat, C. F., Li, S. C., Li, Z., Usaj, M. M., Okada, H., Pascoe, N., San Luis, B.-J., Sharifpoor, S., Shuteriqi, E., Simpkins, S. W., Snider, J., Suresh, H. G., Tan, Y., Zhu, H., Malod-Dognin, N., Janjic, V., Przulj, N., Troyanskaya, O. G., Stagljar, I., Xia, T., Ohya, Y., Gingras, A.-C., Raught, B., Boutros, M., Steinmetz, L. M., Moore, C. L., Rosebrock, A. P., Caudy, A. A., Myers, C. L., Andrews, B., & Boone, C. (2016). A global genetic interaction network maps a wiring diagram of cellular function. *Science*, **353**(6306). doi:10.1126/science.aaf1420

Costas, C., De La Paz Sanchez, M., Stroud, H., Yu, Y., Oliveros, J. C., Feng, S., Benguria, A., López-Vidriero, I., Zhang, X., Solano, R., Jacobsen, S. E., & Gutierrez, C. (2011). Genome-wide mapping of Arabidopsis thaliana origins of DNA replication and their associated epigenetic marks. *Nature Structural and Molecular Biology*, **18**(3), 395–400.

Cotobal, C., Segurado, M., & Antequera, F. (2010). Structural diversity and dynamics of genomic replication origins in Schizosaccharomyces pombe. *EMBO Journal*, **29**(5), 934–942.

Cottarel, G., Shero, J. H., Hieter, P., & Hegemann, J. H. (1989). A 125 bp CEN6 DNA fragment is sufficient for complete meiotic and mitotic centromere functions in Saccharomyces cerevisiae. *Trends in Genetics*, **9**(8), 3342–3349.

Courbet, S., Gay, S., Arnoult, N., Wronka, G., Anglana, M., Brison, O., & Debatisse, M. (2008). Replication fork movement sets chromatin loop size and origin choice in mammalian cells. *Nature*, **455**(7212), 557–560.

Coverley, D., Pelizon, C., Trewick, S., & Laskey, R. A. (2000). Chromatin-bound Cdc6

persists in S and G2 phases in human cells, while soluble Cdc6 is destroyed in a cyclin A-cdk2 dependent process. *Journal of Cell Science*, **113**(11), 1929–1938.

Cullen, J. K., Hussey, S. P., Walker, C., Prudden, J., Wee, B.-Y., Dave, A., Findlay, J. S., Savory, A. P., & Humphrey, T. C. (2007). Break-Induced Loss of Heterozygosity in Fission Yeast: Dual Roles for Homologous Recombination in Promoting Translocations and Preventing De Novo Telomere Addition. *Molecular and Cellular Biology*, **27**(21), 7745–7757.

Cullmann, G., Fien, K., Kobayashi, R., & Stillman, B. (1995). Characterization of the five replication factor C genes of Saccharomyces cerevisiae. *Molecular and Cellular Biology*, **15**(9), 4661–4671.

Dai, J., Chuang, R. Y., & Kelly, T. J. (2005). DNA replication origins in the Schizosaccharomyces pombe genome. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(2), 337–342.

Dalgaard, J. Z., Eydmann, T., Koulintchenko, M., Sayrac, S., Vengrova, S., & Yamada-Inagawa, T. (2009). Random and site-specific replication termination. *Methods in Molecular Biology*, **521**, 35–53.

Das, S. P., Borrman, T., Liu, V. W. T., Yang, S. C. H., Bechhoefer, J., & Rhind, N. (2015). Replication timing is regulated by the number of MCMs loaded at origins. *Genome Research*, **25**(12), 1886–1892.

Dasgupta, S., Masukata, H., & Tomizawa, J. (1987). Multiple mechanisms for initiation of ColE1 DNA replication: DNA synthesis in the presence and absence of ribonuclease H. *Cell*, **51**(6), 1113–1122.

Davé, A., Cooley, C., Garg, M., & Bianchi, A. (2014). Protein Phosphatase 1 recruitment by Rif1 regulates DNA replication origin firing by counteracting DDK activity. *Cell Reports*, **7**(1), 53–61.

Davey, M. J., Fang, L., McInerney, P., Geórgescu, R. E., & O'Donnell, M. (2002). The DnaC helicase loader is a dual ATP/ADP switch protein. *EMBO Journal*, **21**(12), 3148–3159.

Davidson, I. F., Li, A., & Blow, J. J. (2006). Deregulated replication licensing causes DNA fragmentation consistent with head-to-tail fork collision. *Molecular Cell*, **24**(3), 433–443.

Davis, A. J., Yan, Z., Martinez, B., & Mumby, M. C. (2008). Protein phosphatase 2A is targeted to cell division control protein 6 by a calcium-binding regulatory subunit. *Journal of Biological Chemistry*, **283**(23), 16104–16114.

Davis, A. P., & Symington, L. S. (2004). RAD51-dependent break-induced replication in yeast. *Molecular & Cellular Biology*, **24**(6), 2344–2351.

de Massy, B., Fayet, O., & Kogoma, T. (1984). Multiple origin usage for DNA replication in sdrA(rnh) mutants of Escherichia coli K-12. Initiation in the absence of oriC. *Journal of Molecular Biology*, **178**(2), 227–236.

Debatisse, M., Le Tallec, B., Letessier, A., Dutrillaux, B., & Brison, O. (2012). Common fragile sites: Mechanisms of instability revisited. *Trends in Genetics*, **28**(1), 22–32.

Deem, A., Keszthelyi, A., Blackgrove, T., Vayl, A., Coffey, B., Mathur, R., Chabes, A., & Malkova, A. (2011). Break-induced replication is highly inaccurate. *PLoS Biology*, **9**(2), e1000594.

Delgado, S., Gómez, M., Bird, A., & Antequera, F. (1998). Initiation of DNA replication at CpG islands in mammalian chromosomes. *EMBO Journal*, **17**(8), 2426–2435.

Delmolino, L. M., Saha, P., & Dutta, A. (2001). Multiple mechanisms regulate subcellular localization of human CDC6. *Journal of Biological Chemistry*, **276**(29), 26947–26954.

DeMoss, J. A. (1965). Biochemical diversity in the tryptophan pathway. *Biophys. Res. Commun*, **18**, 850–857.

Deog Su Hwang, Thony, B., & Kornberg, A. (1992). IciA protein, a specific inhibitor of initiation of Escherichia coli chromosomal replication. *Journal of Biological Chemistry*, **267**(4), 2209–2213.

DePamphilis, M. L. (2003). The “ORC cycle”: A novel pathway for regulating eukaryotic DNA replication. *Gene*, **310**, 1–15.

Dershowitz, A., Snyder, M., Sbia, M., Skurnick, J. H., Ong, L. Y., & Newlon, C. S. (2007). Linear derivatives of Saccharomyces cerevisiae chromosome III can be maintained in the absence of autonomously replicating sequence elements. *Molecular and Cellular Biology*, **27**(13), 4652–4663.

Descorps-Declère, S., Saguez, C., Cournac, A., Marbouty, M., Rolland, T., Ma, L., Bouchier, C., Moszer, I., Dujon, B., Koszul, R., & Richard, G. F. (2015). Genome-wide replication landscape of Candida glabrata. *BMC Biology*, **13**, 69.

Deshpande, A. M., & Newlon, C. S. (1992). The ARS consensus sequence is required for chromosomal origin function in Saccharomyces cerevisiae. *Molecular and Cellular Biology*, **12**(10), 4305–4313.

Desprat, R., Thierry-Mieg, D., Lailler, N., Lajugie, J., Schildkraut, C., Thierry-Mieg, J., & Bouhassira, E. E. (2009). Predictable dynamic program of timing of DNA replication in human cells. *Genome Research*, **19**(12), 2288–2299.

Devbhandari, S., Jiang, J., Kumar, C., Whitehouse, I., & Remus, D. (2017). Chromatin Constrains the Initiation and Elongation of DNA Replication. *Molecular Cell*, **65**(1), 131–141.

Dewar, J. M., Budzowska, M., & Walter, J. C. (2015). The mechanism of DNA replication termination in vertebrates. *Nature*, **525**(7569), 345–350.

Dewar, J. M., Low, E., Mann, M., Räschle, M., & Walter, J. C. (2017). CRL2Lrr1 promotes unloading of the vertebrate replisome from chromatin during replication termination. *Genes and Development*, **31**(3), 275–290.

Dewar, J. M., & Walter, J. C. (2017). Mechanisms of DNA replication termination. *Nature Reviews Molecular Cell Biology*, **18**(8), 507–516.

Dieckman, L. M., Johnson, R. E., Prakash, S., & Washington, M. T. (2010). Pre-steady state kinetic studies of the fidelity of nucleotide incorporation by yeast DNA polymerase δ. *Biochemistry*, **49**(34), 7344–7350.

Diffley, J. F. X., & Cocker, J. H. (1992). Protein-DNA interactions at a yeast replication origin. *Nature*, **357**(6374), 169–172.

Diffley, J. F. X., Cocker, J. H., Dowell, S. J., Harwood, J., & Rowley, A. (1995). Stepwise assembly of initiation complexes at budding yeast replication origins during the cell cycle. *Journal of Cell Science*, **19**, 67–72.

Diffley, J. F. X., Cocker, J. H., Dowell, S. J., & Rowley, A. (1994). Two steps in the assembly of complexes at yeast replication origins in vivo. *Cell*, **78**(2), 303–316.

Dilley, R. L., Verma, P., Cho, N. W., Winters, H. D., Wondisford, A. R., & Greenberg, R. A. (2016). Break-induced telomere synthesis underlies alternative telomere maintenance. *Nature*, **539**(7627), 54–58.

Dimude, J. U., Stockum, A., Midgley-Smith, S. L., Upton, A. L., Foster, H. A., Khan, A., Saunders, N. J., Retkute, R., & Rudolph, C. J. (2015). The consequences of replicating in the wrong orientation: Bacterial chromosome duplication without an active replication origin. *MBio*, **6**(6), 1–13.

DiNardo, S., Voelkel, K., & Sternglanz, R. (1984). DNA topoisomerase II mutant of Saccharomyces cerevisiae: Topoisomerase II is required for segregation of daughter molecules at the termination of DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*, **81**(9), 2616–2620.

Dion, M. F., Kaplan, T., Kim, M., Buratowski, S., Friedman, N., & Rando, O. J. (2007). Dynamics of replication-independent histone turnover in budding yeast. *Science*, **315**(5817), 1405–1408.

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., & Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*, **485**(7398), 376–380.

Donaldson, A. D. (2005). Shaping time: Chromatin structure and the DNA replication programme. *Trends in Genetics*, **21**(8), 444–449.

Donnianni, R. A., & Symington, L. S. (2013). Break-induced replication occurs by conservative DNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(33), 13475–13480.

Dorn, E. S., & Cook, J. G. (2011). Nucleosomes in the neighborhood: New roles for

chromatin modifications in replication origin control. *Epigenetics*, **6**(5), 552–529.

Douglas, M. E., Ali, F. A., Costa, A., & Diffley, J. F. X. (2018). The mechanism of eukaryotic CMG helicase activation. *Nature*, **555**(7695), 265–268.

Dua, R., Levy, D. L., & Campbell, J. L. (1999). Analysis of the essential functions of the C-terminal protein/protein interaction domain of Saccharomyces cerevisiae pol ε and its unexpected ability to support growth in the absence of the DNA polymerase domain. *Journal of Biological Chemistry*, **274**(32), 22283–22288.

Dubey, D. D., Zhu, J., Carlson, D. L., Sharma, K., & Huberman, J. A. (1994). Three ARS elements contribute to the ura4 replication origin region in the fission yeast, Schizosaccharomyces pombe. *The EMBO Journal*, **13**(15), 3638–3647.

Dueber, E. L., Corn, J. E., Bell, S. D., & Berger, J. M. (2007). Replication origin recognition and deformation by a heterodimeric archaeal Orc1 complex. *Science*, **317**(5842), 1210–1213.

Dungrawala, H., Rose, K. L., Bhat, K. P., Mohni, K. N., Glick, G. G., Couch, F. B., & Cortez, D. (2015). The Replication Checkpoint Prevents Two Types of Fork Collapse without Regulating Replisome Stability. *Molecular Cell*, **59**(6), 998–1010.

Duxin, J. P., Dewar, J. M., Yardimci, H., & Walter, J. C. (2014). Repair of a DNA-protein crosslink by replication-coupled proteolysis. *Cell*, **159**(2), 346–357.

Eaton, M. L., Galani, K., Kang, S., Bell, S. P., & MacAlpine, D. M. (2010). Conserved nucleosome positioning defines replication origins. *Genes and Development*, **24**(8), 748–753.

Eaton, M. L., Prinz, J. A., MacAlpine, H. K., Tretyakov, G., Kharchenko, P. V., & MacAlpine, D. M. (2011). Chromatin signatures of the Drosophila replication program. *Genome Research*, **21**(2), 164–174.

Eder, P. S., & Walder, J. A. (1991). Ribonuclease H from K562 human erythroleukemia cells: Purification, characterization, and substrate specificity. *Journal of Biological Chemistry*, **266**(10), 6472–6479.

Edwards, S., Li, C. M., Levy, D. L., Brown, J., Snow, P. M., & Campbell, J. L. (2003).

Saccharomyces cerevisiae DNA polymerase epsilon and polymerase sigma interact physically and functionally, suggesting a role for polymerase epsilon in sister chromatid cohesion. *Molecular and Cellular Biology*, **23**(8), 2733–2748.

Ehmsen, K. T., & Heyer, W. D. (2008). Saccharomyces cerevisiae Mus81-Mms4 is a catalytic, DNA structure-selective endonuclease. *Nucleic Acids Research*, **36**(7), 2182–2195.

El Hage, A., Webb, S., Kerr, A., & Tollervey, D. (2014). Genome-Wide Distribution of RNA-DNA Hybrids Identifies RNase H Targets in tRNA Genes, Retrotransposons and Mitochondria. *PLoS Genetics*, **10**(10), e1004716.

Elble, R., & Tye, B. K. (1992). Chromosome loss, hyperrecombination, and cell cycle arrest in a yeast mcm1 mutant. *Molecular Biology of the Cell*, **3**(9), 971–980.

Elserafy, M., & El-Khamisy, S. F. (2018). Choose your yeast strain carefully: The RAD5 gene matters. *Nature Reviews Molecular Cell Biology*, **19**(6), 343–344.

Evans, D. R. H., & Stark, M. J. R. (1997). Mutations in the Saccharomyces cerevisiae type 2A protein phosphatase catalytic subunit reveal roles in cell wall integrity, actin cytoskeleton organization and mitosis. *Genetics*, **145**(2), 227–241.

Fachinetti, D., Bermejo, R., Cocito, A., Minardi, S., Katou, Y., Kanoh, Y., Shirahige, K., Azvolinsky, A., Zakian, V. A., & Foiani, M. (2010). Replication Termination at Eukaryotic Chromosomes Is Mediated by Top2 and Occurs at Genomic Loci Containing Pausing Elements. *Molecular Cell*, **39**(4), 595–605.

Fang, D., Lengronne, A., Shi, D., Forey, R., Skrzypczak, M., Ginalski, K., Yan, C., Wang, X., Cao, Q., Pasero, P., & Lou, H. (2017). Dbf4 recruitment by forkhead transcription factors defines an upstream rate-limiting step in determining origin firing timing. *Genes and Development*, **31**(23–24), 2405–2415.

Farkash-Amar, S., Lipson, D., Polten, A., Goren, A., Helmstetter, C., Yakhini, Z., & Simon, I. (2008). Global organization of replication time zones of the mouse genome. *Genome Research*, **18**(10), 1562–1570.

Farkash-Amar, S., & Simon, I. (2010). Genome-wide analysis of the replication program

in mammals. *Chromosome Research*, **18**(1), 115–125.

Fasching, C. L., Cejka, P., Kowalczykowski, S. C., & Heyer, W.-D. (2015). Top3-Rmi1 dissolve Rad51-mediated D-loops by a topoisomerase-based mechanism. *Mol Cell*, **57**(4), 595–606.

Feng, W., Collingwood, D., Boeck, M. E., Fox, L. A., Alvino, G. M., Fangman, W. L., Raghuraman, M. K., & Brewer, B. J. (2006). Genomic mapping of single-stranded DNA in hydroxyurea-challenged yeasts identifies origins of replication. *Nature Cell Biology*, **8**(2), 148–155.

Feng, W., & D'Urso, G. (2001). Schizosaccharomyces pombe Cells Lacking the Amino-Terminal Catalytic Domains of DNA Polymerase Epsilon Are Viable but Require the DNA Damage Checkpoint Control. *Molecular and Cellular Biology*, **21**(14), 4495–4504.

Fenouil, R., Cauchy, P., Koch, F., Descostes, N., Cabeza, J. Z., Innocenti, C., Ferrier, P., Spicuglia, S., Gut, M., Gut, I., & Andrau, J. C. (2012). CpG islands and GC content dictate nucleosome depletion in a transcription-independent manner at mammalian promoters. *Genome Research*, **22**(12), 2399–2408.

Fernandez-Sarabia, M. J., Sutton, A., Zhong, T., & Arndt, K. T. (1992). SIT4 protein phosphatase is required for the normal accumulation of SWI4, CLN1, CLN2, and HCS26 RNAs during late G1. *Genes and Development*, **6**(12A), 2417–2428.

Ferreira, J., & Carmo-Fonseca, M. (1997). Genome replication in early mouse embryos follows a defined temporal and spatial order. *Journal of Cell Science*, **110**(7), 889–897.

Finn, K. J., & Li, J. J. (2013). Single-stranded annealing induced by re-initiation of replication origins provides a novel and efficient mechanism for generating copy number expansion via non-allelic homologous recombination. *PLoS Genetics*, **9**(1), e1003192.

Forsburg, S. L., & Nurse, P. (1991). Cell cycle regulation in the yeasts Saccharomyces cerevisiae and Schizosaccharomyces pombe. *Annual Review of Cell Biology*, **7**,

227–256.

Foss, E. J., Lao, U., Dalrymple, E., Adrianse, R. L., Loe, T., & Bedalov, A. (2017). *SIR2* suppresses replication gaps and genome instability by balancing replication between repetitive and unique sequences. *Proceedings of the National Academy of Sciences*, **114**(3), 552–557.

Foss, M., McNally, F. J., Laurenson, P., & Rine, J. (1993). Origin recognition complex (ORC) in transcriptional silencing and DNA replication in S. cerevisiae. *Science*, **262**(5141), 1838–1844.

Foti, R., Gnan, S., Cornacchia, D., Dileep, V., Bulut-Karslioglu, A., Diehl, S., Buness, A., Klein, F. A., Huber, W., Johnstone, E., Loos, R., Bertone, P., Gilbert, D. M., Manke, T., Jenuwein, T., & Buonomo, S. C. B. (2016). Nuclear architecture organized by Rif1 underpins the replication-timing program. *Molecular Cell*, **61**(2), 260–273.

Foury, F. (1997). Human genetic diseases: A cross-talk between man and yeast. *Gene*, **195**(1), 1–10.

Foury, F., Roganti, T., Lecrenier, N., & Purnelle, B. (1998). The complete sequence of the mitochondrial genome of Saccharomyces cerevisiae. *FEBS Letters*, **440**(3), 325–331.

Frank, P., Braunshofer-Reiter, C., Wintersberger, U., Grimm, R., & Büsen, W. (1998). Cloning of the cDNA encoding the large subunit of human RNase HI, a homologue of the prokaryotic RNase HII. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 12872–12877.

Franz, A., Orth, M., Pirson, P. A., Sonneville, R., Blow, J. J., Gartner, A., Stemmann, O., & Hoppe, T. (2011). CDC-48/p97 Coordinates CDT-1 Degradation with GINS Chromatin Dissociation to Ensure Faithful DNA Replication. *Molecular Cell*, **44**(1), 85–96.

Friedman, K. L., Brewer, B. J., Fangman, W. L., & Yamamoto, M. (1997). Replication profile of Saccharomyces cerevisiae chromosome VI. *Genes to Cells*, **2**(11), 667–678.

Friedman, K. L., Diller, J. D., Ferguson, B. M., Nyland, S. V. M., Brewer, B. J., & Fangman, W. L. (1996). Multiple determinants controlling activation of yeast replication origins late in S phase. *Genes and Development*, **10**(13), 1595–1607.

Fu, Y. V., Yardimci, H., Long, D. T., Ho, V., Guainazzi, A., Bermudez, V. P., Hurwitz, J., Van Oijen, A., Schärer, O. D., & Walter, J. C. (2011). Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell*, **146**(6), 931–941.

Fullbright, G., Rycenga, H. B., Gruber, J. D., & Long, D. T. (2016). p97 Promotes a Conserved Mechanism of Helicase Unloading during DNA Cross-Link Repair. *Molecular and Cellular Biology*, **36**(23), 2983–2994.

Furukawa, T., Ishibashi, T., Kimura, S., Tanaka, H., Hashimoto, J., & Sakaguchi, K. (2003). Characterization of all the subunits of replication factor C from a higher plant, rice (Oryza sativa L.), and their relation to development. *Plant Molecular Biology*, **53**((1-2)), 15–25.

Futcher, A. B. (1988). The 2 μm circle plasmid of Saccharomyces cerevisiae. *Yeast*, **70**(1), 2–17.

Gagou, M. E., Ganesh, A., Phear, G., Robinson, D., Petermann, E., Cox, A., & Meuth, M. (2014). Human PIF1 helicase supports DNA replication and cell growth under oncogenic-stress. *Oncotarget*, **5**(22), 11381–11398.

Gambus, A., Jones, R. C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R. D., & Labib, K. (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nature Cell Biology*, **8**(4), 358–366.

Gambus, A., Van Deursen, F., Polychronopoulos, D., Foltman, M., Jones, R. C., Edmondson, R. D., Calzada, A., & Labib, K. (2009). A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase α within the eukaryotic replisome. *EMBO Journal*, **28**(19), 2992–3004.

Gao, F., & Zhang, C. T. (2007). DoriC: A database of oriC regions in bacterial genomes. *Bioinformatics*, **23**(14), 1866–1867.

Gao, F., & Zhang, C. T. (2008). Ori-Finder: A web-based system for finding oriCs in unannotated bacterial genomes. *BMC Bioinformatics*, **9**, 79.

Gardner, R., Putnam, C. W., & Weinert, T. (1999). RAD53, DUN1 and PDS1 define two parallel G2/M checkpoint pathways in budding yeast. *EMBO Journal*, **18**(11), 3173–3185.

Gaudier, M., Schuwirth, B. S., Westcott, S. L., & Wigley, D. B. (2007). Structural basis of DNA replication origin recognition by an ORC protein. *Science*, **24**(3), 316–324.

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., Arkin, A. P., Astromoff, A., El Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K. D., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein, M., Gotte, D., Güldener, U., Hegemann, J. H., Hempel, S., Herman, Z., Jaramillo, D. F., Kelly, D. E., Kelly, S. L., Kötter, P., LaBonte, D., Lamb, D. C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S. L., Revuelta, J. L., Roberts, C. J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D. D., Sookhai-Mahadeo, S., Storms, R. K., Strathern, J. N., Valle, G., Voet, M., Volckaert, G., Wang, C. yun, Ward, T. R., Wilhelmy, J., Winzeler, E. A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J. D., Snyder, M., Philippsen, P., Davis, R. W., & Johnston, M. (2002). Functional profiling of the Saccharomyces cerevisiae genome. *Nature*, **418**(6896), 387–391.

Giaever, G., & Nislow, C. (2014). The yeast deletion collection: A decade of functional genomics. *Genetics*, **197**(2), 451–465.

Gillespie, P. J., & Blow, J. J. (2010). Clusters, factories and domains: The complex structure of S-phase comes into focus. *Cell Cycle*, **9**(16), 3218–3226.

Gineau, L., Cognet, C., Kara, N., Lach, F. P., Dunne, J., Veturi, U., Picard, C., Trouillet, C., Eidenschenk, C., Aoufouchi, S., Alcaïs, A., Smith, O., Geissmann, F., Feighery, C., Abel, L., Smogorzewska, A., Stillman, B., Vivier, E., Casanova, J. L., & Jouanguy, E. (2012). Partial MCM4 deficiency in patients with growth retardation,

adrenal insufficiency, and natural killer cell deficiency. *Journal of Clinical Investigation*, **122**(3), 821–832.

Goffeau, A., Barrell, G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., & Oliver, S. G. (1996). Life with 6000 genes. *Science*, **274**(5287), 563–567.

Goh, P. Y., Hong, H. L., & Surana, U. (2000). Cdc20 protein contains a destruction-box but, unlike Clb2, its proteolysis is not acutely dependent on the activity of anaphase-promoting complex. *European Journal of Biochemistry*, **267**(2), 434–449.

Göksenin, A. Y., Zahurancik, W., LeCompte, K. G., Taggart, D. J., Suo, Z., & Pursell, Z. F. (2012). Human DNA polymerase ε is able to efficiently extend from multiple consecutive ribonucleotides. *Journal of Biological Chemistry*, **287**(51), 42675–42684.

Goodman, M. F., & Woodgate, R. (2013). Translesion DNA polymerases. *Cold Spring Harbor Perspectives in Biology*, **5**(10), a010363.

Goren, A., Tabib, A., Hecht, M., & Cedar, H. (2008). DNA replication timing of the human β-globin domain is controlled by histone modification at the origin. *Genes and Development*, **22**(10), 1319–1324.

Gowrishankar, J. (2015). End of the beginning: Elongation and termination features of alternative modes of chromosomal replication initiation in bacteria. *PLoS Genetics*, **11**(1), e1004909.

Grainge, I., Gaudier, M., Schuwirth, B. S., Westcott, S. L., Sandall, J., Atanassova, N., & Wigley, D. B. (2006). Biochemical analysis of a DNA replication origin in the archaeon aeropyrum pernix. *Journal of Molecular Biology*, **363**(2), 355–369.

Grallert, A., Boke, E., Hagting, A., Hodgson, B., Connolly, Y., Griffiths, J. R., Smith, D. L., Pines, J., & Hagan, I. M. (2015). A PP1-PP2A phosphatase relay controls mitotic progression. *Nature*, **517**(7532), 94–98.

Gravel, S., Chapman, J. R., Magill, C., & Jackson, S. P. (2008). DNA helicases Sgs1 and

BLM promote DNA double-strand break resection. *Genes and Development*, **22**(20), 2767–2772.

Green, B. M., Finn, K. J., & Li, J. J. (2010). Loss of DNA replication control is a potent inducer of gene amplification. *Science*, **329**(5994), 943–946.

Greenfeder, S. A., & Newlon, C. S. (1992). A replication map of a 61-kb circular derivative of Saccharomyces cerevisiae chromosome III. *Molecular Biology of the Cell*, **3**(9), 999–1013.

Gros, J., Devbhandari, S., & Remus, D. (2014). Origin plasticity during budding yeast DNA replication in vitro. *The EMBO Journal*, **33**(6), 621–636.

Gros, J., Kumar, C., Lynch, G., Yadav, T., Whitehouse, I., & Remus, D. (2015). Post-licensing specification of eukaryotic replication origins by facilitated Mcm2-7 sliding along DNA. *Mol Cell*, **60**(8), 797–807.

Guernsey, D. L., Matsuoka, M., Jiang, H., Evans, S., MacGillivray, C., Nightingale, M., Perry, S., Ferguson, M., Leblanc, M., Paquette, J., Patry, L., Rideout, A. L., Thomas, A., Orr, A., McMaster, C. R., Michaud, J. L., Deal, C., Langlois, S., Superneau, D. W., Parkash, S., Ludman, M., Skidmore, D. L., & Samuels, M. E. (2011). Mutations in origin recognition complex gene ORC4 cause Meier-Gorlin syndrome. *Nature Genetics*, **43**(4), 360–364.

Guillou, E., Ibarra, A., Coulon, V., Casado-Vela, J., Rico, D., Casal, I., Schwob, E., Losada, A., & Méndez, J. (2010). Cohesin organizes chromatin loops at DNA replication factories. *Genes and Development*, **24**(24), 2812–2822.

Hamdan, S. M., Loparo, J. J., Takahashi, M., Richardson, C. C., & Van Oijen, A. M. (2009). Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature*, **457**(7227), 336–369.

Hamlin, J. L., Mesner, L. D., Lar, O., Torres, R., Chodaparambil, S. V., & Wang, L. (2008). A revisionist replicon model for higher eukaryotic genomes. *Journal of Cellular Biochemistry*, **105**(2), 321–329.

Hamperl, S., & Cimprich, K. A. (2016). Conflict resolution in the genome: How

transcription and replication make it work. *Cell*, **167**(6), 1455–1467.

Hardy, C. F. S., Dryga, O., Seematter, S., Pahl, P. M. B., & Sclafani, R. A. (1997). mcm5/cdc46-bob1 bypasses the requirement for the S phase activator Cdc7p. *Proceedings of the National Academy of Sciences of the United States of America*, **94**(7), 3151–3155.

Hartman, A. L., Norais, C., Badger, J. H., Delmas, S., Haldenby, S., Madupu, R., Robinson, J., Khouri, H., Ren, Q., Lowe, T. M., Maupin-Furlow, J., Pohlschroder, M., Daniels, C., Pfeiffer, F., Allers, T., & Eisen, J. A. (2010). The complete genome sequence of Haloferax volcanii DS2, a model archaeon. *PLoS ONE*, **5**(3), e9605.

Hartwell, L. H. (1974). Saccharomyces cerevisiae cell cycle. *Bacteriological Reviews*, **38**(2), 164–198.

Hartwell, L. H., Hood, L., Goldberg, M. L., Reynolds, A. E., Silver, L. M., & Veres, R. C. (2004). Saccharomyces cerevisiae: genetic portrait of a yeast. In L. H. Hartwell, L. Hood, M. L. Goldberg, A. E. Reynolds, L. M. Silver, & R. C. Veres, eds., *Genetics: From Genes to Genomes*, second, United States: McGraw-Hill Education, pp. 739–753.

Hartwell, L. H., Mortimer, R. K., Culotti, J., & Culotti, M. (1973). Genetic control of the cell division cycle in yeast: V. Genetic analysis of cdc mutants. *Genetics*, **74**(2), 267–286.

Hawkins, M., Malla, S., Blythe, M. J., & Nieduszynski, C. A. (2013a). Accelerated growth in the absence of DNA replication origins. *Nature*, **503**(7477), 544–547.

Hawkins, M., Retkute, R., Müller, C. A., Saner, N., Tanaka, T. U., DeMoura, A. P. S., & Nieduszynski, C. A. (2013b). High-resolution replication profiles define the stochastic nature of genome replication initiation and termination. *Cell Reports*, **5**(4), 1132–1141.

Hayashi, M., Katou, Y., Itoh, T., Tazumi, M., Yamada, Y., Takahashi, T., Nakagawa, T., Shirahige, K., & Masukata, H. (2007). Genome-wide localization of pre-RC sites and identification of replication origins in fission yeast. *EMBO Journal*, **26**(5), 1327–

1339.

Hediger, F., Dubrana, K., & Gasser, S. M. (2002). Myosin-like proteins 1 and 2 are not required for silencing or telomere anchoring, but act in the Tel1 pathway of telomere length control. *Journal of Structural Biology*, **40**(1–3), 79–91.

Heller, R. C., Kang, S., Lam, W. M., Chen, S., Chan, C. S., & Bell, S. P. (2011). Eukaryotic origin-dependent DNA replication in vitro reveals sequential action of DDK and S-CDK kinases. *Cell*, **146**(1), 80–91.

Hemerly, A. S., Prasanth, S. G., Siddiqui, K., & Stillman, B. (2009). Orc1 controls centriole and centrosome copy number in human cells. *Science*, **323**(5915), 789–793.

Hiraga, S. I., Alvino, G. M., Chang, F., Lian, H. Y., Sridhar, A., Kubota, T., Brewer, B. J., Weinreich, M., Raghuraman, M. K., & Donaldson, A. D. (2014). Rif1 controls DNA replication by directing Protein Phosphatase 1 to reverse Cdc7- mediated phosphorylation of the MCM complex. *Genes and Development*, **28**(4), 372–383.

Hiraga, S. I., Robertson, E. D., & Donaldson, A. D. (2006). The Ctf18 RFC-like complex positions yeast telomeres but does not specify their replication time. *EMBO Journal*, **25**(7), 1505–1514.

Hiraga, S., Ly, T., Garzón, J., Hořejší, Z., Ohkubo, Y., Endo, A., Obuse, C., Boulton, S. J., Lamond, A. I., & Anne, D. (2016). Human RIF1 and Protein Phosphatase 1 stimulate DNA replication origin licensing but suppress origin activation. *EMBO Reports*, **18**(3), 403–419.

Hiraiwa, A., Fujita, M., Nagasaka, T., Adachi, A., Ohashi, M., & Ishibashi, M. (1997). Immunolocalization of the hCDC47 protein in normal and neoplastic human tissues and its relation to growth. *International Journal of Cancer*, **74**(2), 180–184.

Hiratani, I., Ryba, T., Itoh, M., Yokochi, T., Schwaiger, M., Chang, C. W., Lyou, Y., Townes, T. M., Schübeler, D., & Gilbert, D. M. (2008). Global reorganization of replication domains during embryonic stem cell differentiation. *PLoS Biology*, **6**(10), e245.

Hoang, M. L., Leon, R. P., Pessoa-Brandao, L., Hunt, S., Raghuraman, M. K., Fangman, W. L., Brewer, B. J., & Sclafani, R. A. (2007). Structural Changes in Mcm5 Protein Bypass Cdc7-Dbf4 Function and Reduce Replication Origin Efficiency in Saccharomyces cerevisiae. *Molecular and Cellular Biology*, **27**(21), 7594–7602.

Hogg, M., Osterman, P., Bylund, G. O., Ganai, R. A., Lundström, E. B., Sauer-Eriksson, A. E., & Johansson, E. (2014). Structural basis for processive DNA synthesis by yeast DNA polymerase É. *Nature Structural and Molecular Biology*, **21**(1), 49–55.

Hohmann, S., Krantz, M., & Nordlander, B. (2007). Yeast Osmoregulation. *Methods in Enzymology*, **428**, 29–45.

Honey, S., & Futcher, B. (2007). Roles of the CDK phosphorylation sites of yeast Cdc6 in chromatin binding and rereplication. *Molecular Biology of the Cell*, **18**(4), 1324–1336.

Hong, X., Cadwell, G. W., & Kogoma, T. (1995). Escherichia coli RecG and RecA proteins in R-loop formation. *The EMBO Journal*, **14**(10), 2385–2392.

Hook, S. S., Lin, J. J., & Dutta, A. (2007). Mechanisms to control rereplication and implications for cancer. *Current Opinion in Cell Biology*, **19**(6), 663–671.

Hughes, C. R., Guasti, L., Meimaridou, E., Chuang, C. H., Schimenti, J. C., King, P. J., Costigan, C., Clark, A. J. L., & Metherell, L. A. (2012a). MCM4 mutation causes adrenal failure, short stature, and natural killer cell deficiency in humans. *Journal of Clinical Investigation*, **122**(3), 814–820.

Hughes, S., Elustondo, F., Di Fonzo, A., Leroux, F. G., Wong, A. C., Snijders, A. P., Matthews, S. J., & Cherepanov, P. (2012b). Crystal structure of human CDC7 kinase in complex with its activator DBF4. *Nature Structural and Molecular Biology*, **19**(11), 1101–1107.

Hyrien, O., Marheineke, K., & Goldar, A. (2003). Paradoxes of eukaryotic DNA replication: MCM proteins and the random completion problem. *BioEssays*, **25**(2), 116–125.

Iizuka, M., Matsui, T., Takisawa, H., & Smith, M. M. (2006). Regulation of replication

licensing by acetyltransferase Hbo1. *Molecular and Cellular Biology*, **26**(3), 1098–1108.

Ilves, I., Petojevic, T., Pesavento, J. J., & Botchan, M. R. (2010). Activation of the MCM2-7 Helicase by Association with Cdc45 and GINS Proteins. *Molecular Cell*, **37**(2), 247–258.

Ira, G., & Haber, J. E. (2002). Characterization of RAD51-independent break-induced replication that acts preferentially with short homologous sequences. *Mol Cell Biol*, **22**(18), 6384–6392.

Ivanova, T., Maier, M., Missarova, A., Ziegler-Birling, C., Carey, L. B., & Mendoza, M. (2018). Budding yeast complete DNA replication after chromosome segregation begins. *BioRxiv*, **407957**. doi:10.1101/407957

Ivessa, A. S., Zhou, J. Q., & Zakian, V. A. (2000). The Saccharomyces Pif1p DNA helicase and the highly related Rrm3p have opposite effects on replication fork progression in ribosomal DNA. *Cell*, **100**(4), 479–489.

Iyer, L. M., Leipe, D. D., Koonin, E. V., & Aravind, L. (2004). Evolutionary history and higher order classification of AAA+ ATPases. *Journal of Structural Biology*, **146**(1–2), 11–31.

Jackson, D. A., & Pombo, A. (1998). Replicon clusters are stable units of chromosome structure: Evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. *Journal of Cell Biology*, **140**(6), 1285–1295.

Jackson, L. P., Reed, S. I., & Haase, S. B. (2006). Distinct Mechanisms Control the Stability of the Related S-Phase Cyclins Clb5 and Clb6. *Molecular and Cellular Biology*, **26**(6), 2456–2466.

Jacob, F., Brenner, S., & Cuzin, F. (1963). On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon. *Cold Spring Harbor Quantitative Biology*, **256**, 298–300.

Jaiswal, R., Choudhury, M., Zaman, S., Singh, S., Santosh, V., Bastia, D., & Escalante,

C. R. (2016). Functional architecture of the Reb1-Ter complex of Schizosaccharomyces pombe. *Proceedings of the National Academy of Sciences of the United States of America*, **113**(16), E2267-76.

Jares, P., Donaldson, A., & Blow, J. J. (2000). The Cdc7/Dbf4 protein kinase: Target of the S phase checkpoint? *EMBO Reports*, **1**(4), 319–322.

Jeon, Y., Bekiranov, S., Karnani, N., Kapranov, P., Ghosh, S., MacAlpine, D., Lee, C., Deog, S. H., Gingeras, T. R., & Dutta, A. (2005). Temporal profile of replication of human chromosomes. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(18), 6419–6424.

Jiang, F., & Doudna, J. A. (2017). CRISPR–Cas9 structures and mechanisms. *Annual Review of Biophysics*, **46**, 505–529.

Jiang, W., McDonald, D., Hope, T. J., & Hunter, T. (1999a). Mammalian Cdc7-Dbf4 protein kinase complex is essential for initiation of DNA replication. *EMBO Journal*, **18**(20), 5703–5713.

Jiang, W., Wells, N. J., & Hunter, T. (1999b). Multistep regulation of DNA replication by Cdk phosphorylation of HsCdc6. *Proceedings of the National Academy of Sciences of the United States of America*, **96**(11), 6193–6198.

Jimenez, J., Bru, S., Ribeiro, M., & Clotet, J. (2015). Live fast, die soon: cell cycle progression and lifespan in yeast cells. *Microbial Cell*, **2**(3), 62–67.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., & Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark “nucleosome-free regions” of active promoters and other regulatory regions. *Nature Genetics*, **41**(8), 941–945.

Johansson, E., & Dixon, N. (2013). Replicative DNA polymerases. *Cold Spring Harbor Perspectives in Biology*, **5**(6), a012799.

Johnson, A., & Skotheim, J. M. (2013). Start and the restriction point. *Current Opinion in Cell Biology*, **25**(6), 717–723.

Johnson, R. E., Klassen, R., Prakash, L., & Prakash, S. (2015). A major role of DNA

polymerase delta in replication of both the leading and lagging DNA strands. *Molecular Cell*, **59**(2), 163–175.

Johnston, L. H., & Barker, D. G. (1987). Characterisation of an autonomously replicating sequence from the fission yeast Schizosaccharomyces pombe. *MGG Molecular & General Genetics*, **207**(1), 161–164.

Johnston, L. H., & Nasmyth, K. A. (1978). Saccharomyces cerevisiae cell cycle mutant cdc9 is defective in DNA ligase. *Nature*, **274**(5674), 891–893.

Johnston, L. H., & Thomas, A. P. (1982). The isolation of new DNA synthesis mutants in the yeast Saccharomyces cerevisiae. *Mol Gen Genet*, **186**(3), 439–444.

Kaguni, J. M. (2011). Replication initiation at the Escherichia coli chromosomal origin. *Current Opinion in Chemical Biology*, **15**(5), 606–613.

Kaliraman, V., Mullen, J. R., Fricke, W. M., Bastin-Shanower, S. A., & Brill, S. J. (2001). Functional overlap between Sgs1-Top3 and the Mms4-Mus81 endonuclease. *Genes and Development*, **15**(20), 2730–2740.

Kamimura, Y., Tak, Y. S., Sugino, A., & Araki, H. (2001). Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in saccharomyces cerevisiae. *EMBO Journal*, **20**(8), 2097–2107.

Kanemaki, M., & Labib, K. (2006). Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. *EMBO Journal*, **25**(8), 1753–1763.

Kang, Y. H., Farina, A., Bermudez, V. P., Tappin, I., Du, F., Galal, W. C., & Hurwitz, J. (2013). Interaction between human Ctf4 and the Cdc45/Mcm2-7/GINS (CMG) replicative helicase. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(49), 19760–19765.

Kang, Y. H., Lee, C. H., & Seo, Y. S. (2010). Dna2 on the road to Okazaki fragment processing and genome stability in eukaryotes. *Critical Reviews in Biochemistry and Molecular Biology*, **45**(2), 71–96.

Kaplan, D. L., Davey, M. J., & O'Donnell, M. (2003). Mcm4,6,7 uses a "pump in ring"

mechanism to unwind DNA by steric exclusion and actively translocate along a duplex. *The Journal of Biological Chemistry*, **278**(49), 49171–49182.

Karathia, H., Vilaprinyo, E., Sorribas, A., & Alves, R. (2011). Saccharomyces cerevisiae as a model organism: A comparative study. *PLoS ONE*, **6**(2), e16015.

Katou, Y., Kanoh, Y., Bando, M., Noguchi, H., Tanaka, H., Ashikari, T., Sugimoto, K., & Shirahige, K. (2003). S-phase checkpoint proteins Tof1 and Mrc1 form a stable replication-pausing complex. *Nature*, **424**(6952), 1078–1083.

Kawabata, T., Luebben, S. W., Yamaguchi, S., Ilves, I., Matise, I., Buske, T., Botchan, M. R., & Shima, N. (2011). Stalled fork rescue via dormant replication origins in unchallenged S phase promotes proper chromosome segregation and tumor suppression. *Molecular Cell*, **41**(5), 543–553.

Kawame, H., Gartler, S. M., & Hansen, R. S. (1995). Allele-specific replication timing in imprinted domains: Absence of asynchrony at several loci. *Human Molecular Genetics*, **4**(12), 2287–2293.

Ke, Y., Xu, Y., Chen, X., Feng, S., Liu, Z., Sun, Y., Yao, X., Li, F., Zhu, W., Gao, L., Chen, H., Du, Z., Xie, W., Xu, X., Huang, X., & Liu, J. (2017). 3D chromatin structures of mature gametes and structural reprogramming during mammalian embryogenesis. *Cell*, **170**(2), 367–381.

Kearsey, S. (1983). Analysis of sequences conferring autonomous replication in bakers's yeast. *Embo J*, **2**(1571), 1571–1575.

Kedziora, S., Gali, V. K., Wilson, R. H. C., Clark, K. R. M., Nieduszynski, C. A., Hiraga, S. I., & Donaldson, A. D. (2018). Rif1 acts through Protein Phosphatase 1 but independent of replication timing to suppress telomere extension in budding yeast. *Nucleic Acids Research*, **46**(8), 3993–4003.

Kesti, T., Flick, K., Keränen, S., Syväoja, J. E., & Wittenberg, C. (1999). DNA polymerase ε catalytic domains are dispensable for DNA replication, DNA repair, and cell viability. *Molecular Cell*, **3**(5), 679–685.

Keszthelyi, A., Daigaku, Y., Ptasin, K., Miyabe, I., Carr, A. M., Ptasińska, K., Miyabe, I.,

& Carr, A. M. (2015). Mapping ribonucleotides in genomic DNA and exploring replication dynamics by polymerase usage sequencing (Pu-seq). *Nature Protocols*, **10**(11), 1786–1801.

Kim, S. M., & Huberman, J. A. (2001). Regulation of replication timing in fission yeast. *EMBO Journal*, **20**(21), 6115–6126.

Kinoshita, N., Yamano, H., Niwa, H., Yoshida, T., & Yanagida, M. (1993). Negative regulation of mitosis by the fission yeast protein phosphatase ppa2. *Genes and Development*, **7**(6), 1059–1071.

Kinoshita, Y., Johnson, E. M., Gordon, R. E., Negri-Bell, H., Evans, M. T., Coolbaugh, J., Rosario-Peralta, Y., Samet, J., Slusser, E., Birkenbach, M. P., & Daniel, D. C. (2008). Colocalization of MCM8 and MCM7 with proteins involved in distinct aspects of DNA replication. *Microscopy Research and Technique*, **71**(4), 288–297.

Kitamura, E., Tanaka, K., Kitamura, Y., & Tanaka, T. U. (2007). Kinetochore-microtubule interaction during S phase in Saccharomyces cerevisiae. *Genes and Development*, **21**(24), 3319–3330.

Kitsberg, D., Selig, S., Brandels, M., Simon, I., Keshet, I., Driscoll, D. J., Nicholls, R. D., & Cedar, H. (1993). Allele-specific replication timing of imprinted gene regions. *Nature*, **364**(6436), 459–463.

Klemm, R. D., & Bell, S. P. (2001). ATP bound to the origin recognition complex is important for preRC formation. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(15), 8361–8367.

Klotz-Noack, K., & Blow, J. J. (2011). A role for dormant origins in tumor suppression. *Molecular Cell*, **41**(5), 495–496.

Knott, S. R. V, Peace, J. M., Ostrow, A. Z., Gan, Y., Rex, A. E., Viggiani, C. J., Tavaré, S., & Aparicio, O. M. (2012). Forkhead transcription factors establish origin timing and long-range clustering in S. cerevisiae. *Cell*, **148**(1–2), 99–111.

Knott, S. R. V, Viggiani, C. J., Tavaré, S., & Aparicio, O. M. (2009). Genome-wide replication profiles indicate an expansive role for Rpd3L in regulating replication

initiation timing or efficiency, and reveal genomic loci of Rpd3 function in Saccharomyces cerevisiae. *Genes and Development*, **23**(9), 1077–1090.

Koepp, D. M., Kile, A. C., Swaminathan, S., & Rodriguez-Rivera, V. (2006). The F-box protein Dia2 regulates DNA replication. *Molecular Biology of the Cell*, **17**(4), 1540–1548.

Kogoma, T. (1978). A novel Escherichia coli mutant capable of DNA replication in the absence of protein synthesis. *Journal of Molecular Biology*, **121**(1), 55–69.

Koh, K. D., Balachander, S., Hesselberth, J. R., & Storici, F. (2015). Ribose-seq: Global mapping of ribonucleotides embedded in genomic DNA. *Nature Methods*, **12**(3), 251–257.

Koren, A., Tsai, H., Tirosh, I., Burrack, L. S., Barkai, N., & Berman, J. (2010). Epigenetically-inherited centromere and neocentromere DNA replicates earliest in s-phase. *PLoS Genetics*, **6**(8), e1001068.

Kowalski, D., & Eddy, M. J. (1989). The DNA unwinding element: a novel, cis-acting component that facilitates opening of the Escherichia coli replication origin. *The EMBO Journal*, **8**(13), 4335–4344.

Kreuzer, K. N. (2000). Recombination-dependent DNA replication in phage T4. *Trends in Biochemical Sciences*, **25**(4), 165–173.

Kubota, T., Nishimura, K., Kanemaki, M. T., & Donaldson, A. D. (2013). The Elg1 replication factor C-like complex functions in PCNA unloading during DNA replication. *Molecular Cell*, **50**(2), 273–280.

Kumagai, A., & Dunphy, W. G. (2000). Claspin, a novel protein required for the activation of Chk1 during a DNA replication checkpoint response in Xenopus egg extracts. *Molecular Cell*, **6**(4), 839–849.

Kumagai, A., Shevchenko, A., Shevchenko, A., & Dunphy, W. G. (2010). Treslin collaborates with TopBP1 in triggering the initiation of DNA replication. *Cell*, **140**(3), 349–359.

Kumagai, A., Shevchenko, A., Shevchenko, A., & Dunphy, W. G. (2011). Direct

regulation of Treslin by cyclin-dependent kinase is essential for the onset of DNA replication. *Journal of Cell Biology*, **193**(6), 995–1007.

Kumaran, R., Yang, S. Y., & Leu, J. Y. (2013). Characterization of chromosome stability in diploid, polyploid and hybrid yeast cells. *PLoS ONE*, **8**(7), e68094.

Kunnev, D., Rusiniak, M. E., Kudla, A., Freeland, A., Cady, G. K., & Pruitt, S. C. (2010). DNA damage response and tumorigenesis in Mcm2-deficient mice. *Oncogene*, **29**(25), 3630–3638.

Kuo, A. J., Song, J., Cheung, P., Ishibe-Murakami, S., Yamazoe, S., Chen, J. K., Patel, D. J., & Gozani, O. (2012). The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature*, **484**(7392). doi:10.1038/nature10956

Kupiec, M. (2014). Biology of telomeres: Lessons from budding yeast. *FEMS Microbiology Reviews*, **38**(2), 144–171.

Kurat, C. F., Yeeles, J. T. P., Patel, H., Early, A., & Diffley, J. F. X. (2017). Article Chromatin Controls DNA Replication Origin Selection , Lagging Strand Synthesis , and Replication Fork Rates. *Molecular Cell*, **65**(1), 117–130.

Kuzminov, A., & Stahl, F. W. (1999). Double-strand end repair via the RecBC pathway in Escherichia coli primes DNA replication Double-strand end repair via the RecBC pathway in Escherichia coli primes DNA replication. *Genes Dev.*, **13**, 345–356.

Labib, K. (2010). How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells? *Genes and Development*, **24**(12), 1208–1219.

Labib, K., Diffley, J. F. X., & Kearsey, S. E. (1999). G1-phase and B-type cyclins exclude the DNA-replication factor Mcm4 from the nucleus. *Nature Cell Biology*, **1**(7), 415–422.

Lebofsky, R., Heilig, R., Sonnleitner, M., Weissenbach, J., & Bensimon, A. (2006). DNA replication origin interference increases the spacing between initiation events in human cells. *Molecular Biology of the Cell*, **17**(12), 5337–5345.

Lee, D. Y., & Clayton, D. A. (1998). Initiation of mitochondrial DNA replication by transcription and R-loop processing. *Journal of Biological Chemistry*, **273**(46), 30614–30621.

Lee, K. Y., Bae, J. S., Yoon, S., & Hwang, D. S. (2014). Dephosphorylation of Orc2 by protein phosphatase 1 promotes the binding of the origin recognition complex to chromatin. *Biochemical and Biophysical Research Communications*, **448**(4), 385–389.

Lee, T. H., Solomon, M. J., Mumby, M. C., & Kirschner, M. W. (1991). INH, a negative regulator of MPF, is a form of protein phosphatase 2A. *Cell*, **64**(2), 415–423.

Lei, M., Kawasaki, Y., Young, M. R., Kihara, M., Sugino, A., & Tye, B. K. (1997). Mcm2 is a target of regulation by Cdc7-Dbf4 during the initiation of DNA synthesis. *Genes and Development*, **11**(24), 3365–3374.

Lei, M., & Tye, B. K. (2001). Initiating DNA synthesis: From recruiting to activating the MCM complex. *Journal of Cell Science*, **114**(8), 1447–1454.

Leonard, A. C., & Mechali, M. (2013). DNA replication origins. *Cold Spring Harbor Perspectives in Medicine*, **5**(10), a010116.

Letessier, A., Millot, G. A., Koundrioukoff, S., Lachagès, A. M., Vogt, N., Hansen, R. S., Malfoy, B., Brison, O., & Debatisse, M. (2011). Cell-type-specific replication initiation programs set fragility of the FRA3B fragile site. *Nature*, **470**(7332), 120–123.

Leung-Pineda, V., Ryan, C. E., & Piwnica-Worms, H. (2006). Phosphorylation of Chk1 by ATR Is antagonized by a Chk1-regulated protein phosphatase 2A circuit. *Molecular and Cellular Biology*, **26**(20), 7529–7538.

Liachko, I., Bhaskar, A., Lee, C., Chung, S. C. C., Tye, B. K., & Keich, U. (2010). A comprehensive genome-wide map of autonomously replicating sequences in a naive genome. *PLoS Genetics*, **6**(5), e1000946.

Liachko, I., Tanaka, E., Cox, K., Chung, S. C. C., Yang, L., Seher, A., Hallas, L., Cha, E., Kang, G., Pace, H., Barrow, J., Inada, M., Tye, B. K., & Keich, U. (2011). Novel features of ARS selection in budding yeast Lachancea kluyveri. *BMC Genomics*,

12, 633.

Lian, H.-Y., Robertson, E. D., Hiraga, S. -i., Alvino, G. M., Collingwood, D., McCune, H. J., Sridhar, A., Brewer, B. J., Raghuraman, M. K., & Donaldson, A. D. (2011). The effect of Ku on telomere replication time is mediated by telomere length but is independent of histone tail acetylation. *Molecular Biology of the Cell*, **22**(10), 1753–1765.

Lianga, N., Williams, E. C., Kennedy, E. K., Doré, C., Pilon, S., Girard, S. L., Deneault, J. S., & Rudner, A. D. (2013). A wee1 checkpoint inhibits anaphase onset. *Journal of Cell Biology*, **201**(6), 843–862.

Lieberman-Aiden, E., Van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., & Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, **326**(5950), 289–293.

Liu, J., Xu, L., Sandler, S. J., & Marians, K. J. (1999). Replication fork assembly at recombination intermediates is required for bacterial growth. *Proceedings of the National Academy of Sciences of the United States of America*, **96**(7), 3552–3555.

Lombraña, R., Almeida, R., Álvarez, A., & Gómez, M. (2015). R-loops and initiation of DNA replication in human cells: A missing link? *Frontiers in Genetics*, **6**, 158.

Lombraña, R., Alvarez, A., Fernandez-Justel, J. M., Almeida, R., Poza-Carrion, C., Gomes, F., Calzada, A., Requena, J. M., & Gomez, M. (2016). Transcriptionally Driven DNA Replication Program of the Human Parasite Leishmania major. *Cell Reports*, **16**(6), 1774–1786.

Long, D. T., Joukov, V., Budzowska, M., & Walter, J. C. (2014). BRCA1 promotes unloading of the CMG Helicase from a stalled DNA replication fork. *Molecular Cell*, **56**(1), 174–185.

Longhese, M. P., Plevani, P., & Lucchini, G. (1994). Replication factor A is required in

vivo for DNA replication, repair, and recombination. *Molecular and Cellular Biology*, **14**(12), 7884–7890.

Longtine, M. S., Enomoto, S., Finstad, S. L., & Berman, J. (1992). Yeast telomere repeat sequence (TRS) improves circular plasmid segregation, and TRS plasmid segregation involves the RAP1 gene product. *Molecular and Cellular Biology*, **12**(5), 1997–2009.

Loog, M., & Morgan, D. O. (2005). Cyclin specificity in the phosphorylation of cyclin-dependent kinase substrates. *Nature*, **434**(7029), 104–108.

López-Mirabal, H. R., & Winther, J. R. (2008). Redox characteristics of the eukaryotic cytosol. *Biochimica et Biophysica Acta - Molecular Cell Research*, **1783**(4), 629–640.

Louis, E. J. (1995). The chromosome ends of Saccharomyces cerevisiae. *Yeast*, **11**(16), 1553–1573.

Louis, E. J., & Haber, J. E. (1990). The subtelomeric Y' repeat family in Saccharomyces cerevisiae: an experimental system for repeated sequence evolution. *Genetics*, **124**(3), 533–545.

Louis, E. J., & Haber, J. E. (1992). The structure and evolution of subtelomeric Y' repeats in Saccharomyces cerevisiae. *Genetics*, **131**(3), 559–574.

Louis, E. J., Naumova, E. S., Lee, A., Naumov, G., & Haber, J. E. (1994). The chromosome end in yeast: Its mosaic nature and influence on recombinational dynamics. *Genetics*, **136**(3), 789–802.

Lucas, I., Germe, T., Chevrier-Miller, M., & Hyrien, O. (2001). Topoisomerase II can unlink replicating DNA by precatenane removal. *EMBO Journal*, **20**(22), 6509–6519.

Lujan, S. A., Williams, J. S., Clausen, A. R., Clark, A. B., & Kunkel, T. A. (2013). Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Molecular Cell*, **50**(3), 437–443.

Lujan, S. A., Williams, J. S., Pursell, Z. F., Abdulovic-Cui, A. A., Clark, A. B., Nick

McElhinny, S. A., & Kunkel, T. A. (2012). Mismatch Repair Balances Leading and Lagging Strand DNA Replication Fidelity. *PLoS Genetics*, **8**(10), e1003016.

Lundgren, M., Andersson, A., Chen, L., Nilsson, P., & Bernander, R. (2004). Three replication origins in Sulfolobus species: Synchronous initiation of chromosome replication and asynchronous termination. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(18), 7046–7051.

Lutzmann, M., & Méchali, M. (2008). MCM9 binds Cdt1 and is required for the assembly of prereplication complexes. *Molecular Cell*, **31**(2), 190–200.

Lydeard, J. R., Jain, S., Yamaguchi, M., & Haber, J. E. (2007). Break-induced replication and telomerase-independent telomere maintenance require Pol32. *Nature*, **448**(7155), 820–823.

MacAlpine, D. M., Rodríguez, H. K., & Bell, S. P. (2004). Coordination of replication and transcription along a Drosophila chromosome. *Genes and Development*, **18**(24), 3094–3105.

MacAlpine, H. K., Gordân, R., Powell, S. K., Hartemink, A. J., & MacAlpine, D. M. (2010). Drosophila ORC localizes to open chromatin and marks sites of cohesin complex loading. *Genome Research*, **20**(2), 201–211.

Mackiewicz, P., Zakrzewska-Czerwińska, J., Zawilak, A., Dudek, M. R., & Cebrat, S. (2004). Where does bacterial replication start? Rules for predicting the oriC region. *Nucleic Acids Research*, **32**(13), 3781–3791.

MacNeill, S. A. (2010). Structure and function of the GINS complex, a key component of the eukaryotic replisome. *Biochemical Journal*, **425**(3), 489–500.

Maduike, N. Z., Tehranchi, A. K., Wang, J. D., & Kreuzer, K. N. (2014). Replication of the Escherichia coli chromosome in RNase HI- deficient cells: Multiple initiation regions and fork dynamics. *Molecular Microbiology*, **91**(1), 39–56.

Majka, J., & Burgers, P. M. J. (2004). The PCNA-RFC families of DNA clamps and clamp loaders. *Progress in Nucleic Acid Research and Molecular Biology*, **78**, 227–260.

Makarova, K. S., & Koonin, E. V. (2013). Archaeology of eukaryotic DNA replication. *Cold*

Spring Harbor Perspectives in Biology, **5**(11), a012963.

Makarova, K. S., Krupovic, M., & Koonin, E. V. (2014). Evolution of replicative DNA polymerases in archaea and their contributions to the eukaryotic replication machinery. *Frontiers in Microbiology*, **5**, 354.

Mantiero, D., MacKenzie, A., Donaldson, A., & Zegerman, P. (2011). Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. *EMBO Journal*, **30**(23), 4805–4814.

Marahrens, Y., & Stillman, B. (1992). A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science*, **255**(5046), 817–823.

Margolis, S. S., Perry, J. A., Weitzel, D. H., Freel, C. D., Yoshida, M., Haystead, T. A., & Kornbluth, S. (2006). A role for PP1 in the Cdc2/cyclin B-mediated positive feedback activation of Cdc25. *Molecular Biology of the Cell*, **17**(4), 1779–1789.

Margolis, S. S., Walsh, S., Weiser, D. C., Yoshida, M., Shenolikar, S., & Kornbluth, S. (2003). PP1 control of M phase entry exerted through 14-3-3-regulated Cdc25 dephosphorylation. *EMBO Journal*, **22**(21), 5734–5745.

Maric, M., Maculins, T., De Piccoli, G., & Labib, K. (2014). Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication. *Science*, **346**(6208), 1253596.

Marszalek, J., & Kaguni, J. M. (1994). DnaA protein directs the binding of DnaB protein in initiation of DNA replication in Escherichia coli. *Journal of Biological Chemistry*, **269**(7), 4883–4890.

Martínez, P., & Blasco, M. A. (2015). Replicating through telomeres: A means to an end. *Trends in Biochemical Sciences*, **40**(9), 504–515.

Masai, H., Taniyama, C., Ogino, K., Matsui, E., Kakusho, N., Matsumoto, S., Kim, J. M., Ishii, A., Tanaka, T., Kobayashi, T., Tamai, K., Ohtani, K., & Arai, K. I. (2006). Phosphorylation of MCM4 by Cdc7 kinase facilitates its interaction with Cdc45 on the chromatin. *Journal of Biological Chemistry*, **281**(51), 39249–39261.

Masukata, H., Dasgupta, S., & Tomizawa, J. ichi. (1987). Transcriptional activation of

ColE1 DNA synthesis by displacement of the nontranscribed strand. *Cell*, **51**(6), 1123–1130.

Masumoto, H., Muramatsu, S., Kamimura, Y., & Araki, H. (2002). S-Cdk-dependent phosphorylation of Sld2 essential for chromosomal DNA replication in budding yeast. *Nature*, **415**(6872), 651–655.

Matheson, K., Parsons, L., & Gammie, A. (2017). Whole-genome sequence and variant analysis of W303, a widely-used strain of Saccharomyces cerevisiae. *G3: Genes, Genomes, Genetics*, **7**(7), 2219–2226.

Matsuno, K., Kumano, M., Kubota, Y., Hashimoto, Y., & Takisawa, H. (2006). The N-Terminal Noncatalytic Region of Xenopus RecQ4 Is Required for Chromatin Binding of DNA Polymerase in the Initiation of DNA Replication. *Molecular and Cellular Biology*, **26**(13), 4843–4852.

Mattarocci, S., Shyian, M., Lemmens, L., Damay, P., Altintas, D. M., Shi, T., Bartholomew, C. R., Thomä, N. H., Hardy, C. F. J., & Shore, D. (2014). Rif1 controls DNA replication timing in yeast through the PP1 Phosphatase Glc7. *Cell Reports*, **7**(1), 62–69.

Maundrell, K., Hutchison, A., & Shall, S. (1988). Sequence analysis of ARS elements in fission yeast. *The EMBO Journal*, **7**(7), 2203–2209.

Mayle, R., Campbell, I. M., Beck, C. R., Yu, Y., Wilson, M., Shaw, C. A., Bjergbaek, L., Lupski, J. R., & Ira, G. (2015). Mus81 and converging forks limit the mutagenicity of replication fork breakage. *Science*, **349**(6249), 742–747.

McCune, H. J., Danielson, L. S., Alvino, G. M., Collingwood, D., Delrow, J. J., Fangman, W. L., Brewer, B. J., & Raghuraman, M. K. (2008). The temporal program of chromosome replication: Genomewide replication in clb5Δ Saccharomyces cerevisiae. *Genetics*, **180**(4), 1833–1847.

McDonald, K. R., Guise, A. J., Pourbozorgi-Langroudi, P., Cristea, I. M., Zakian, V. A., Capra, J. A., & Sabouri, N. (2016). Pfh1 Is an Accessory Replicative Helicase that Interacts with the Replisome to Facilitate Fork Progression and Preserve Genome

Integrity. *PLoS Genetics*, **12**(9), e1006238.

McEachern, M. J., & Blackburn, E. H. (1994). A conserved sequence motif within the exceptionally diverse telomeric sequences of budding yeasts. *Proceedings of the National Academy of Sciences of the United States of America*, **91**(8), 3453–3457.

McEachern, M. J., & Hicks, J. B. (1993). Unusually large telomeric repeats in the yeast Candida albicans. *Molecular and Cellular Biology*, **13**(1), 551–560.

McElhinny, N. S., Gordenin, D., Stith, C., Burgers, P., & Kunkel, T. (2008). Division of labor at the eukaryotic replication fork. *Mol Cell*, **30**(2), 137–144.

McElhinny, S. A. N., Kissling, G. E., & Kunkel, T. A. (2010a). Differential correction of lagging-strand replication errors made by DNA polymerases α and δ. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(49), 21070–21075.

McElhinny, S. A. N., Kumar, D., Clark, A. B., Watt, D. L., Watts, B. E., Lundström, E. B., Johansson, E., Chabes, A., & Kunkel, T. A. (2010b). Genome instability due to ribonucleotide incorporation into DNA. *Nature Chemical Biology*, **6**(10), 774–781.

McGuffee, S. R., Smith, D. J., & Whitehouse, I. (2013). Quantitative, Genome-Wide Analysis of Eukaryotic Replication Initiation and Termination. *Molecular Cell*, **50**(1), 123–135.

Méchali, M. (2010). Eukaryotic DNA replication origins: Many choices for appropriate answers. *Nature Reviews Molecular Cell Biology*, **11**(10), 728–738.

Mendez, J., & Stillman, B. (2000). Chromatin Association of Human Origin Recognition Complex, Cdc6, and Minichromosome Maintenance Proteins during the Cell Cycle: Assembly of Prereplication Complexes in Late Mitosis. *Molecular and Cellular Biology*, **20**(22), 8602–8612.

Merlet, J., Burger, J., Tavernier, N., Richaudeau, B., Gomes, J. E., & Pintard, L. (2010). The CRL2LRR-1 ubiquitin ligase regulates cell cycle progression during C. elegans development. *Development*, **137**(22), 3857–3866.

Miller-Fleming, L., Giorgini, F., & Outeiro, T. F. (2008). Yeast as a model for studying

human neurodegenerative disorders. *Biotechnology Journal*, **3**(3), 325–338.

Miotto, B., & Struhl, K. (2008). HBO1 histone acetylase is a coactivator of the replication licensing factor Cdt1. *Genes and Development*, **22**(19), 2633–2638.

Miotto, B., & Struhl, K. (2010). HBO1 histone acetylase activity is essential for DNA replication licensing and inhibited by geminin. *Molecular Cell*, **37**(1), 57–66.

Miyabe, I., Kunkel, T. A., & Carr, A. M. (2011). The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. *PLoS Genetics*, **7**(12), e1002407.

Miyabe, I., Mizuno, K., Keszthelyi, A., Daigaku, Y., Skouteri, M., Mohebi, S., Kunkel, T. A., Murray, J. M., & Carr, A. M. (2015). Polymerase δ replicates both strands after homologous recombination–dependent fork restart. *Nature Structural & Molecular Biology*, **22**(11), 932–938.

Mizushima, T., Takahashi, N., & Stillman, B. (2000). Cdc6p modulates the structure and DNA binding activity of the origin recognition complex in vitro. *Genes and Development*, **14**(13), 1631–1641.

Mocciaro, A., & Schiebel, E. (2010). Cdc14: A highly conserved family of phosphatases with non-conserved functions? *Journal of Cell Science*, **123**(17), 2867–2876.

Mohanty, B. K., Bairwa, N. K., & Bastia, D. (2006). The Tof1p-Csm3p protein complex counteracts the Rrm3p helicase to control replication termination of Saccharomyces cerevisiae. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(4), 897–902.

Moindrot, B., Audit, B., Klous, P., Baker, A., Thermes, C., De Laat, W., Bouvet, P., Mongelard, F., & Arneodo, A. (2012). 3D chromatin conformation correlates with replication timing and is conserved in resting cells. *Nucleic Acids Research*, **40**(19), 9470–9481.

Moir, D., Stewart, S. E., Osmond, B. C., & Botstein, D. (1982). Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties, and pseudoreversion studies. *Genetics*, **100**(4), 547–563.

Moreno, S. P., Bailey, R., Campion, N., Herron, S., & Gambus, A. (2014). Polyubiquitylation drives replisome disassembly at the termination of DNA replication. *Science*, **346**(6208), 477–481.

Morgan, D. O. (1997). Cyclin-dependent kinases: Engines, clocks, and microprocessors. *Annual Review of Cell and Developmental Biology*, **13**, 261–291.

Morrison, A., Araki, H., Clark, A. B., Hamatake, R. K., & Sugino, A. (1990). A third essential DNA polymerase in S. cerevisiae. *Cell*, **62**(6), 1143–1151.

Morrow, D. M., Connelly, C., & Hieter, P. (1997). "Break copy" duplication: A model for chromosome fragment formation in saccharomyces cerevisiae. *Genetics*, **147**(2), 371–382.

Moyer, S. E., Lewis, P. W., & Botchan, M. R. (2006). Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(27), 10236–10241.

Müller, C. (2012). *Comparative genomics of chromosome replication in sensu strico yeasts*, University of Nottingham.

Müller, C. A., Boemo, M. A., Spingardi, P., Kessler, B. M., Kriaucionis, S., Simpson, J. T., & Nieduszynski, C. A. (2019). Capturing the dynamics of genome replication on individual ultra-long nanopore sequence reads. *Nature Methods*, **16**(5), 429–436.

Müller, C. A., Hawkins, M., Retkute, R., Malla, S., Wilson, R., Blythe, M. J., Nakato, R., Komata, M., Shirahige, K., De Moura, A. P. S. S., & Nieduszynski, C. A. (2013). The dynamics of genome replication using deep sequencing. *Nucleic Acids Research*, **42**(1), e3.

Müller, C. A., & Nieduszynski, C. A. (2012). Conservation of replication timing reveals global and local regulation of replication origin activity. *Genome Research*, **22**(10), 1953–1962.

Müller, C. A., & Nieduszynski, C. A. (2017). DNA replication timing influences gene expression level. *Journal of Cell Biology*, **216**(7), 1907–1914.

Muller, E. G. D. (1991). Thioredoxin deficiency in yeast prolongs S phase and shortens the G1 interval of the cell cycle. *Journal of Biological Chemistry*, **266**(14), 9194–9202.

Müller, P., Park, S., Shor, E., Huebert, D. J., Warren, C. L., Ansari, A. Z., Weinreich, M., Eaton, M. L., MacAlpine, D. M., & Fox, C. A. (2010). The conserved bromo-adjacent homology domain of yeast Orc1 functions in the selection of DNA replication origins within chromatin. *Genes and Development*, **24**(13), 1418–1433.

Murakami, C., & Kaeberlein, M. (2009). Quantifying yeast chronological life span by outgrowth of aged cells. *Journal of Visualized Experiments*, (27), e1156.

Muramatsu, S., Hirai, K., Tak, Y. S., Kamimura, Y., & Araki, H. (2010). CDK-dependent complex formation between replication proteins Dpb11, Sld2, Pol ε, and GINS in budding yeast. *Genes and Development*, **24**(6), 602–612.

Myllykallio, H., Lopez, P., López-García, P., Heilig, R., Saurin, W., Zivanovic, Y., Philippe, H., & Forterre, P. (2000). Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science*, **288**(5474), 2212–2215.

Nakamura, H., Morita, T., & Sato, C. (1986). Structural organizations of replicon domains during DNA synthetic phase in the mammalian nucleus. *Experimental Cell Research*, **165**(2), 291–297.

Nasheuer, H. P., Smith, R., Bauerschmidt, C., Grosse, F., & Weisshart, K. (2002). Initiation of eukaryotic DNA replication: Regulation and mechanisms. *Progress in Nucleic Acid Research and Molecular Biology*, **72**, 41–94.

Natale, D. A., Schubert, A. E., & Kowalski, D. (1992). DNA helical stability accounts for mutational defects in a yeast replication origin. *Proceedings of the National Academy of Sciences of the United States of America*, **89**(7), 2654–2658.

Natsume, T., Müller, C. A., Katou, Y., Retkute, R., Gierliński, M., Araki, H., Blow, J. J., Shirahige, K., Nieduszynski, C. A., & Tanaka, T. U. (2013). Kinetochores coordinate pericentromeric cohesion and early DNA replication by Cdc7-Dbf4 kinase recruitment. *Molecular Cell*, **50**(5), 661–674.

Neelsen, K. J., Zanini, I. M. Y., Mijic, S., Herrador, R., Zellweger, R., Chaudhuri, A. R., Creavin, K. D., Julian Blow, J., & Lopes, M. (2013). Deregulated origin licensing leads to chromosomal breaks by rereplication of a gapped DNA template. *Genes and Development*, **27**(23), 2537–2542.

Nethanel, T., Reisfeld, S., Dinter-Gottlieb, G., & Kaufmann, G. (1988). An Okazaki piece of simian virus 40 may be synthesized by ligation of shorter precursor chains. *Journal of Virology*, **62**(8), 2867–2873.

Newlon, C. S., Lipchitz, L. R., Collins, I., Deshpande, A., Devenish, R. J., Green, R. P., Klein, H. L., Palzkill, T. G., Ren, R., Synn, S., & Woody, S. T. (1991). Analysis of a circular derivative of Saccharomyces cerevisiae chromosome III: A physical map and identification and location of ARS elements. *Genetics*, **129**(2), 343–357.

Newlon, C. S., & Theis, J. F. (1993). The structure and function of yeast ARS elements. *Current Opinion in Genetics and Development*, **3**(5), 752–758.

Newman, T. J., Mamun, M. A., Nieduszynski, C. A., & Blow, J. J. (2013). Replisome stall events have shaped the distribution of replication origins in the genomes of yeasts. *Nucleic Acids Research*, **41**(21), 9705–9718.

Nguyen, M. O., Jalan, M., Morrow, C. A., Osman, F., & Whitby, M. C. (2015). Recombination occurs within minutes of replication blockage by RTS1 producing restarted forks that are prone to collapse. *ELife*, **2015**(4), 1–25.

Nguyen, V., Co, C., & Li, J. (2001). Cyclin-dependent kinases prevent DNA re-replication through multiple mechanisms. *Nature*, **411**(6841), 1068–1073.

Nguyen, V. Q., Co, C., Irie, K., & Li, J. J. (2000). Clb/Cdc28 kinases promote nuclear export of the replication initiator proteins Mcm2-7. *Current Biology*, **10**(4), 195–205.

Nick McElhinny, S. A., Watts, B. E., Kumar, D., Watt, D. L., Lundström, E. B., Burgers, P. M. J., Johansson, E., Chabes, A., & Kunkel, T. A. (2010). Abundant ribonucleotide incorporation into DNA by yeast replicative polymerases. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(11), 4949–4954.

Nieduszynski, C. A., & Donaldson, A. D. (2009). Detection of replication origins using comparative genomics and recombinational ARS assay. *Methods in Molecular Biology*, **521**, 295–313.

Nieduszynski, C. A., Hiraga, S. I., Ak, P., Benham, C. J., & Donaldson, A. D. (2007). OriDB: A DNA replication origin database. *Nucleic Acids Research*, **35**, D40-6.

Nieduszynski, C. A., Knox, Y., Donaldson, A. D., Nieduszynski, C. A., Knox, Y., & Donaldson, A. D. (2006). Genome-wide identification of replication origins in yeast by comparative genomics, **20**(14), 1874–1879.

Nievera, C., Torgue, J. J. C., Grimwade, J. E. E., & Leonard, A. C. (2006). SeqA blocking of DnaA-oriC interactions ensures staged sssembly of the E. coli pre-RC. *Molecular Cell*, **24**(4), 581–592.

Nishitani, H., & Lygerou, Z. (2002). Control of DNA replication licensing in a cell cycle. *Genes to Cells : Devoted to Molecular & Cellular Mechanisms*, **7**(6), 523–534.

Nishitani, H., Lygerou, Z., & Nishimoto, T. (2004). Proteolysis of DNA replication licensing factor Cdt1 in S-phase is performed independently of Geminin through its N-terminal region. *Journal of Biological Chemistry*, **279**(29), 30807–30816.

Noble, W., Duan, Z. jun, Andronescu, M., Schutz, K., McIlwain, S., Kim, Y. J., Lee, C., Shendure, J., Fields, S., & Blau, C. A. (2011). A Three-Dimensional Model of the Yeast Genome. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, **465**(7296), 363–367.

Noguchi, K., Vassilev, A., Ghosh, S., Yates, J. L., & DePamphilis, M. L. (2006). The BAH domain facilitates the ability of human Orc1 protein to activate replication origins in vivo. *EMBO Journal*, **25**(22), 5372–5382.

Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., Van Berkum, N. L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Blüthgen, N., Dekker, J., & Heard, E. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, **485**(7398), 381–385.

Norais, C., Hawkins, M., Hartman, A. L., Eisen, J. A., Myllykallio, H., & Allers, T. (2007). Genetic and physical mapping of DNA replication origins in Haloferax volcanii. *PLoS Genetics*, **3**(5), 729–743.

Norbury, C. (1992). Animal cell cycles and their control. *Annual Review of Biochemistry*, **61**, 441–470.

Nougarede, R., Della Seta, F., Zarzov, P., & Schwob, E. (2000). Hierarchy of S-phase-promoting factors: Yeast Dbf4-Cdc7 kinase requires prior S-phase cyclin-dependent kinase activation. *Molecular and Cellular Biology*, **20**(11), 3795–3806.

O'Donnell, M., Langston, L., & Stillman, B. (2013). Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harbor Perspectives in Biology*, **5**(7), a010108.

O'Keefe, R. T., Henderson, S. C., & Spector, D. L. (1992). Dynamic organization of DNA replication in mammalian cell nuclei: Spatially and temporally defined replication of chromosome-specific α- satellite DNA sequences. *Journal of Cell Biology*, **116**(5), 1095–1110.

Ogawa, T., Pickett, G. G., Kogoma, T., & Kornberg, A. (1984). RNase H confers specificity in the dnaA-dependent initiation of replication at the unique origin of the Escherichia coli chromosome in vivo and in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, **81**(Feb), 1040–1044.

Ohya, T., Kawasaki, Y., Hiraga, S. I., Kanbara, S., Nakajo, K., Nakashima, N., Suzuki, A., & Sugino, A. (2002). The DNA polymerase domain of polε is required for rapid, efficient, and highly accurate chromosomal DNA replication, telomere length maintenance, and normal cell senescence in Saccharomyces cerevisiae. *Journal of Biological Chemistry*, **277**(31), 28099–28108.

Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K., Kainuma, R., Sugino, A., & Iwatsuki, N. (1968). In vivo mechanism of DNA chain growth. *Cold Spring Harbor Symposia on Quantitative Biology*, **33**, 129–143.

Okuno, Y., Satoh, H., Sekiguchi, M., & Masukata, H. (1999). Clustered adenine/thymine

stretches are essential for function of a fission yeast replication origin. *Molecular and Cellular Biology*, **19**(10), 6699–6709.

Oldach, P., & Nieduszynski, C. A. (2019). Cohesin-mediated genome architecture does not define DNA replication timing domains. *Genes*, **10**(3), 196.

On, K. F., Beuron, F., Frith, D., Snijders, A. P., Morris, E. P., & Diffley, J. F. X. (2014). Prereplicative complexes assembled in vitro support origin-dependent and independent DNA replication. *EMBO Journal*, **33**(6), 605–620.

Orr, S., Gaymes, T., Ladon, D., Chronis1, C., Czepulkowski, B., Wang, R., Mufti, G., EM, Marcotte, & Thomas, N. (2010). Reducing MCM levels in human primary T cells during the G0→G1 transition causes genomic instability during the first cell cycle. *NIH Public Access*, **100**(2), 130–134.

Owsianowski, E., Walter, D., & Fahrenkrog, B. (2008). Negative regulation of apoptosis in yeast. *Biochimica et Biophysica Acta - Molecular Cell Research*, **1783**(7), 1303–1310.

Pacek, M., & Walter, J. C. (2004). A requirement for MCM7 and Cdc45 in chromosome unwinding during eukaryotic DNA replication. *EMBO Journal*, **23**(18), 3667–3676.

Padovan-Merhar, O., Nair, G. P., Biaesch, A. G., Mayer, A., Scarfone, S., Foley, S. W., Wu, A. R., Churchman, L. S., Singh, A., & Raj, A. (2015). Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. *Molecular Cell*, **58**(2), 339–352.

Pak, D. T. S., Pflumm, M., Chesnokov, I., Huang, D. W., Kellum, R., Marr, J., Romanowski, P., & Botchan, M. R. (1997). Association of the origin recognition complex with heterochromatin and HP1 in higher eukaryotes. *Cell*, **91**(3), 311–323.

Pal, G., Paraz, M. T. Z., & Kellogg, D. R. (2008). Regulation of Mih1/Cdc25 by protein phosphatase 2A and casein kinase 1. *Journal of Cell Biology*, **180**(5), 931–945.

Palzkill, T. G., Oliver, S. G., & Newlon, C. S. (1986). DNA sequence analysis of ARS elements from chromosome III of Saccharomyces cerevisiae: Identification of a new conserved sequence. *Nucleic Acids Research*, **14**(15), 6247–6264.

Palzkill, T., & Newlon, C. S. (1988). Yeast replication origin consists of multiple copies of a small conserved sequence. *Cell*, **53**, 441–450.

Pan, X., Ye, P., Yuan, D. S., Wang, X., Bader, J. S., & Boeke, J. D. (2006). A DNA integrity network in the yeast Saccharomyces cerevisiae. *Cell*, **124**(5), 1069–1081.

Pardee, A. B., Dubrow, R., Hamlin, J. L., & Kletzien, R. F. (1978). Animal cell cycle. *Annual Review of Biochemistry*, **47**, 715–750.

Pasion, S. G., & Forsburg, S. L. (1999). Nuclear localization of Schizosaccharomyces pombe Mcm2/Cdc19p requires MCM complex assembly. *Molecular Biology of the Cell*, **10**(12), 4043–4057.

Pellegrini, L. (2012). The Pol α-Primase Complex. *Subcellular Biochemistry*, **62**, 157–169.

Pelve, E. A., Martens-Habbena, W., Stahl, D. A., & Bernander, R. (2013). Mapping of active replication origins in vivo in thaum- and euryarchaeal replicons. *Molecular Microbiology*, **90**(3), 538–550.

Peplowska, K., Wallek, A. U., & Storchova, Z. (2014). Sgo1 regulates both condensin and Ipl1/Aurora B to promote chromosome biorientation. *PLoS Genetics*, **10**(6), e1004411.

Perera, R. L., Torella, R., Klinge, S., Kilkenny, M. L., Maman, J. D., & Pellegrini, L. (2013). Mechanism for priming DNA synthesis by yeast DNA Polymerase α. *ELife*, **2**, e00482.

Petersen, B. O., Lukas, J., Sørensen, C. S., Bartek, J., & Helin, K. (1999). Phosphorylation of mammalian CDC6 by cyclin A/CDK2 regulates its subcellular localization. *EMBO Journal*, **18**(2), 396–410.

Petersen, B. O., Wagener, C., Marinoni, F., Kramer, E. R., Melixetian, M., Denchi, E. L., Gieffers, C., Matteucci, C., Peters, J. M., & Helin, K. (2000). Cell cycle- and cell growth-regulated proteolysis of mammalian CDC6 is dependent on APC-CDH1. *Genes and Development*, **14**(18), 2330–2343.

Picard, A., Capony, J. P., Brautigan, D. L., & Doree, M. (1989). Involvement of protein

phosphatases 1 and 2A in the control of M phase-promoting factor activity in starfish. *Journal of Cell Biology*, **109**(6 Pt 2), 3347–3354.

Pietrobon, V., Fréon, K., Hardy, J., Costes, A., Iraqui, I., Ochsenbein, F., & Lambert, S. A. E. (2014). The chromatin assembly factor 1 promotes Rad51-dependent template switches at replication forks by counteracting D-loop disassembly by the RecQ-type helicase Rqh1. *PLoS Biology*, **12**(10), e1001968.

Poh, W. T., Chadha, G. S., Gillespie, P. J., Kaldis, P., & Blow, J. J. (2014). Xenopus Cdc7 executes its essential function early in S phase and is counteracted by checkpoint-regulated protein phosphatase 1. *Open Biology*, **4**, 130138.

Pohl, T. J., Brewer, B. J., & Raghuraman, M. K. (2012). Functional centromeres determine the activation time of pericentric origins of DNA replication in Saccharomyces cerevisiae. *PLoS Genetics*, **8**(5), e1002677.

Polaczek, P., Kwan, K., Liberles, D. A., & Campbell, J. L. (1997). Role of architectural elements in combinatorial regulation of initiation of DNA replication in Escherichia coli. *Molecular Microbiology*, **26**(2), 261–275.

Pollok, S., & Grosse, F. (2007). Cdc45 degradation during differentiation and apoptosis. *Biochemical and Biophysical Research Communications*, **362**(4), 910–915.

Poloumienko, A., Dershowitz, A., De, J., & Newlon, C. S. (2001). Completion of replication map of Saccharomyces cerevisiae chromosome III. *Molecular Biology of the Cell*, **12**(11), 3317–3327.

Pope, B. D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D. L., Wang, Y., Hansen, R. S., Canfield, T. K., Thurman, R. E., Cheng, Y., Gülsoy, G., Dennis, J. H., Snyder, M. P., Stamatoyannopoulos, J. A., Taylor, J., Hardison, R. C., Kahveci, T., Ren, B., & Gilbert, D. M. (2014). Topologically associating domains are stable units of replication-timing regulation. *Nature*, **515**(7527), 402–405.

Postow, L., Crisona, N. J., Peter, B. J., Hardy, C. D., & Cozzarelli, N. R. (2001). Topological challenges to DNA replication: Conformations at the fork. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(15), 8219–

8226.

Prakash, S., Johnson, R. E., & Prakash, L. (2005). Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annual Review of Biochemistry*, **74**, 317–353.

Prasanth, S. G., Prasanth, K. V., Siddiqui, K., Spector, D. L., & Stillman, B. (2004). Human Orc2 localizes to centrosomes, centromeres and heterochromatin during chromosome inheritance. *EMBO Journal*, **23**(13), 2651–2663.

Pruitt, S. C., Bailey, K. J., & Freeland, A. (2007). Reduced Mcm2 expression results in severe stem/progenitor cell deficiency and cancer. *Stem Cells*, **25**(12), 3121–3132.

Pursell, Z. F., Isoz, I., Lundström, E. B., Johansson, E., & Kunkel, T. A. (2007). Yeast DNA polymerase ε participates in leading-strand DNA replication. *Science*, **317**(5834), 127–130.

Quon, K. C., Yang, B., Domian, I. J., Shapiro, L., & Marczynski, G. T. (1998). Negative control of bacterial DNA replication by a cell cycle regulatory protein that binds at the chromosome origin. *Proceedings of the National Academy of Sciences of the United States of America*, **95**(1), 120–125.

Raghuraman, M., Brewer, B., & Fangman, W. (1997). Cell cycle-dependent establishment of a late replication program. *Science*, **276**(5313), 806–809.

Raghuraman, M. K., Winzeler, E. A., Collingwood, D., Hunt, S., Wodicka, L., Conway, A., Lockhart, D. J., Davis, R. W., Brewer, B. J., & Fangman, W. L. (2001). Replication dynamics of the yeast genome. *Science*, **294**(5540), 115–121.

Ramos, F., Villoria, M. T., Alonso-Rodríguez, E., & Clemente-Blanco, A. (2019). Role of protein phosphatases PP1, PP2A, PP4 and Cdc14 in the DNA damage response. *Cell Stress*, **3**(3), 70–85.

Ravoitytė, B., & Wellinger, R. (2017). Non-Canonical Replication Initiation: You're Fired! *Genes*, **8**(2), 54.

Raymann, K., Forterre, P., Brochier-Armanet, C., & Gribaldo, S. (2014). Global phylogenomic analysis disentangles the complex evolutionary history of DNA

replication in Archaea. *Genome Biology and Evolution*, **6**(1), 192–212.

Reijns, M. A. M., Kemp, H., Ding, J., De Procé, S. M., Jackson, A. P., & Taylor, M. S. (2015). Lagging-strand replication shapes the mutational landscape of the genome. *Nature*, **518**(7540), 502–506.

Remus, D., Beall, E. L., & Botchan, M. R. (2004). DNA topology, not DNA sequence, is a critical determinant for Drosophila ORC-DNA binding. *EMBO Journal*, **23**(4), 897–907.

Remus, D., Beuron, F., Tolun, G., Griffith, J. D., Morris, E. P., & Diffley, J. F. X. X. (2009). Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell*, **139**(4), 719–730.

Rhiannon, B., & T. Young, E. (2009). The disorderly study of ordered recruitment. *Yeast*, **26**(4), 205–220.

Rhind, N., & Gilbert, D. M. (2013). DNA replication timing. *Cold Spring Harbor Perspectives in Medicine*, **5**(8), a010132.

Rippe, K., & Luke, B. (2015). TERRA and the state of the telomere. *Nature Structural & Molecular Biology*, **22**(11), 853–858.

Rivera-Mulia, J. C., & Gilbert, D. M. (2016). Replicating large genomes: Divide and conquer. *Molecular Cell*, **62**(5), 756–765.

Rivier, D. H., & Rine, J. (1992). An origin of DNA replication and a transcription silencer require a common element. *Science*, **256**(5057), 659–663.

Robinson, N. P., & Bell, S. D. (2007). Extrachromosomal element capture and the evolution of multiple replication origins in archaeal chromosomes. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(14), 5806–5811.

Robinson, N. P., Dionne, I., Lundgren, M., Marsh, V. L., Bernander, R., & Bell, S. D. (2004). Identification of two origins of replication in the single chromosome of the archaeon Sulfolobus solfataricus. *Cell*, **116**(1), 25–38.

Roumelioti, F., Sotiriou, S. K., Katsini, V., Chiourea, M., Halazonetis, T. D., & Gagos, S.

(2016). Alternative lengthening of human telomeres is a conservative DNA replication process with features of break-induced replication. *EMBO Reports*, **17**(12), 1731–1737.

Rozgaja, T. A., Grimwade, J. E., Iqbal, M., Czerwonka, C., Vora, M., & Leonard, A. C. (2011). Two oppositely oriented arrays of low-affinity recognition sites in oriC guide progressive binding of DnaA during Escherichiacoli pre-RC assembly. *Molecular Microbiology*, **17**(12), 1731–1737.

Rudolph, C. J., Upton, A. L., Stockum, A., Nieduszynski, C. A., & Lloyd, R. G. (2013). Avoiding chromosome pathology when replication forks collide. *Nature*, **500**(7464), 608–611.

Ryan, V. T., Grimwade, J. E., Camara, J. E., Crooke, E., & Leonard, A. C. (2004). Escherichla coli prereplication complex assembly is regulated by dynamic interplay among Fis, IHF and DnaA. *Molecular Microbiology*, **51**(5), 1347–1359.

Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., Schulz, T. C., Robins, A. J., Dalton, S., & Gilbert, D. M. (2010). Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Research*, **20**(6), 761–770.

Ryba, T., Hiratani, I., Sasaki, T., Battaglia, D., Kulik, M., Zhang, J., Dalton, S., & Gilbert, D. M. (2011). Replication timing: A fingerprint for cell identity and pluripotency. *PLoS Computational Biology*, **7**(10), e1002225.

Saini, N., Ramakrishnan, S., Elango, R., Ayyar, S., Zhang, Y., Deem, A., Ira, G., Haber, J. E., Lobachev, K. S., & Malkova, A. (2013). Migrating bubble during break-induced replication drives conservative DNA synthesis, **502**(7471), 389–392.

Salsi, V., Ferrari, S., Ferraresi, R., Cossarizza, A., Grande, A., & Zappavigna, V. (2009). HOXD13 binds DNA replication origins to promote origin licensing and is inhibited by geminin. *Molecular and Cellular Biology*, **29**(21), 5775–5788.

Samel, S. A., Fernández-Cid, A., Sun, J., Riera, A., Tognetti, S., Carmen Herrera, M., Li, H., & Speck, C. (2014). A unique DNA entry gate serves for regulated loading of the

eukaryotic replicative helicase MCM2-7 onto DNA. *Genes and Development*, **28**(15), 1653–1666.

Samson, R. Y., Xu, Y., Gadelha, C., Stone, T. A., Faqiri, J. N., Li, D., Qin, N., Pu, F., Liang, Y. X., She, Q., & Bell, S. D. (2013). Specificity and function of archaeal DNA replication initiator proteins. *Cell Reports*, **3**(2), 485–496.

Sangrithi, M. N., Bernal, J. A., Madine, M., Philpott, A., Lee, J., Dunphy, W. G., & Venkitaraman, A. R. (2005). Initiation of DNA replication requires the RECQL4 protein mutated in Rothmund-Thomson syndrome. *Cell*, **121**(6), 887–898.

Sansam, C. L., Cruz, N. M., Danielian, P. S., Amsterdam, A., Lau, M. L., Hopkins, N., & Lees, J. A. (2010). A vertebrate gene, ticrr, is an essential checkpoint and replication regulator. *Genes and Development*, **24**(2), 183–194.

Santocanale, C., & Diffley, J. F. (1996). ORC- and Cdc6-dependent complexes at active and inactive chromosomal replication origins in Saccharomyces cerevisiae. *The EMBO Journal*, **15**(23), 6671–6679.

Santocanale, C., Sharma, K., & Diffley, J. F. X. (1999). Activation of dormant origins of DNA replication in budding yeast. *Genes and Development*. doi:10.1101/gad.13.18.2360

Sato, N., Arai, K. I., & Masai, H. (1997). Human and Xenopus cDNAs encoding budding yeast Cdc7-related kinases: In vitro phosphorylation of MCM subunits by a putative human homologue of Cdc7. *EMBO Journal*, **16**(14), 4340–4351.

Saxena, S., & Dutta, A. (2005). Geminin-Cdt1 balance is critical for genetic stability. *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, **569**(1–2), 111–121.

Schacherer, J., Ruderfer, D. M., Gresham, D., Dolinski, K., Botstein, D., & Kruglyak, L. (2007). Genome-wide analysis of nucleotide-level variation in commonly used Saccharomyces cerevisiae strains. *PLoS ONE*, **2**(3), e322.

Schalbetter, S. A., Mansoubi, S., Chambers, A. L., Downs, J. A., & Baxter, J. (2015). Fork rotation and DNA precatenation are restricted during DNA replication to

prevent chromosomal instability. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(33), E4565-70.

Schaper, S., & Messer, W. (1995). Interaction of the initiator protein DnaA of Escherichia coli with its DNA target. *Journal of Biological Chemistry*, **270**(29), 17622–17626.

Schlesinger, S., Selig, S., Bergman, Y., & Cedar, H. (2009). Allelic inactivation of rDNA loci. *Genes and Development*, **23**(20), 2437–2447.

Schübeler, D., Scalzo, D., Kooperberg, C., van Steensel, B., Delrow, J., & Groudine, M. (2002). Genome-wide DNA replication profile for Drosophila melanogaster: A link between transcription and replication timing. *Nature Genetics*, **32**(3), 438–442.

Schwartz, M. F., Duong, J. K., Sun, Z., Morrow, J. S., Pradhan, D., & Stern, D. F. (2002). Rad9 phosphorylation sites couple Rad53 to the Saccharomyces cerevisiae DNA damage checkpoint. *Molecular Cell*, **9**(5), 1055–1065.

Segurado, M., de Luis, A., & Antequera, F. (2003). Genome-wide distribution of DNA replication origins at A+T-rich islands in Schizosaccharomyces pombe. *EMBO Reports*, **4**(11), 1048–1053.

Sengupta, S., Van Deursen, F., De Piccoli, G., & Labib, K. (2013). Dpb2 integrates the leading-strand DNA polymerase into the eukaryotic replisome. *Current Biology*, **23**(7), 543–552.

Sequeira-Mendes, J., Díaz-Uriarte, R., Apedaile, A., Huntley, D., Brockdorff, N., & Gómez, M. (2009). Transcription initiation activity sets replication origin efficiency in mammalian cells. *PLoS Genetics*, **5**(4), e1000446.

Sfeir, A. J., Chai, W., Shay, J. W., & Wright, W. E. (2005). Telomere-end processing: The terminal nucleotides of human chromosomes. *Molecular Cell*, **18**(1), 131–138.

Shaw, S. L., Yeh, E., Maddox, P., Salmon, E. D., & Bloom, K. (1997). Astral microtubule dynamics in yeast: A microtubule-based searching mechanism for spindle orientation and nuclear migration into the bud. *Journal of Cell Biology*, **139**(4), 985–994.

Sherman, F. (2002). Getting started with yeast. In C. Guthrie & G. R. Fink, eds., *Methods*

in Enzymology, Vol 350, Elsevier Science, pp. 3–41.

Sherr, C. J., & Roberts, J. M. (2004). Living with or without cyclins and cyclin-dependent kinases. *Genes and Development*, **18**(22), 2699–2711.

Sheu, Y. J., & Stillman, B. (2006). Cdc7-Dbf4 phosphorylates MCM proteins via a docking site-mediated mechanism to promote S phase progression. *Molecular Cell*, **24**(1), 101–113.

Sheu, Y. J., & Stillman, B. (2010). The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature*, **463**(7277), 113–117.

Shima, N., Alcaraz, A., Liachko, I., Buske, T. R., Andrews, C. A., Munroe, R. J., Hartford, S. A., Tye, B. K., & Schimenti, J. C. (2007). A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. *Nature Genetics*, **39**(1), 93–98.

Shimada, K., Pasero, P., & Gasser, S. M. (2002). ORC and the intra-S-phase checkpoint: A threshold regulates Rad53p activation in S phase. *Genes and Development*, **16**(24), 3236–3252.

Shirahige, K., Iwasaki, T., Rashid, M. B., Ogasawara, N., & Yoshikawa, H. (1993). Location and characterization of autonomously replicating sequences from chromosome VI of Saccharomyces cerevisiae. *Molecular and Cellular Biology*, **13**(8), 5043–5056.

Shu, Y., & Hallberg, R. L. (1995). SCS1, a multicopy suppressor of hsp60-ts mutant alleles, does not encode a mitochondrially targeted protein. *Molecular and Cellular Biology*, **15**(10), 5618–5626.

Simon, A. C., Zhou, J. C., Perera, R. L., Van Deursen, F., Evrin, C., Ivanova, M. E., Kilkenny, M. L., Renault, L., Kjaer, S., Matak-Vinkovi, D., Labib, K., Costa, A., & Pellegrini, L. (2014). A Ctf4 trimer couples the CMG helicase to DNA polymerase α in the eukaryotic replisome. *Nature*, **510**(7504), 293–297.

Simpson, R. T. (1990). Nucleosome positioning can affect the function of a cis-acting DNA elementin vivo. *Nature*, **343**(6256), 387–389.

Sinha, N. K., Morris, C. F., & Alberts, B. M. (1980). Efficient in vitro replication of double-stranded DNA templates by a purified T4 bacteriophage replication system. *Journal of Biological Chemistry*, **255**(9), 4290–4293.

Siow, C. C., Nieduszynska, S. R., Müller, C. A., & Nieduszynski, C. A. (2012). OriDB, the DNA replication origin database updated and extended. *Nucleic Acids Research*, **40**, D682-6.

Slater, S., Wold, S., Lu, M., Boye, E., Skarstad, K., & Kleckner, N. (1995). E. coli SeqA protein binds oriC in two different methyl-modulated reactions appropriate to its roles in DNA replication initiation and origin sequestration. *Cell*, **82**(6), 927–936.

Smith, C. E., Llorente, B., & Symington, L. S. (2007). Template switching during break-induced replication. *Nature*, **447**(7140), 102–105.

Smith, D. J., & Whitehouse, I. (2012). Intrinsic coupling of lagging-strand synthesis to chromatin assembly. *Nature*, **483**(7390), 434–8.

Solovei, I., Kreysing, M., Lanctôt, C., Kösem, S., Peichl, L., Cremer, T., Guck, J., & Joffe, B. (2009). Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. *Cell*, **137**(2), 356–368.

Sonneville, R., Moreno, S. P., Knebel, A., Johnson, C., Hastie, C. J., Gartner, A., Gambus, A., & Labib, K. (2017). CUL-2LRR-1 and UBXN-3 drive replisome disassembly during DNA replication termination and mitosis. *Nature Cell Biology*, **19**(5), 468–479.

Soudet, J., Jolivet, P., & Teixeira, M. T. (2014). Elucidation of the DNA end-replication problem in saccharomyces cerevisiae. *Molecular Cell*, **53**(6), 954–964.

Sparks, J. L., Chon, H., Cerritelli, S. M., Kunkel, T. A., Johansson, E., Crouch, R. J., & Burgers, P. M. (2012). RNase H2-initiated ribonucleotide excision repair. *Molecular Cell*, **47**(6), 980–986.

Speck, C., Chen, Z., Li, H., & Stillman, B. (2005). ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nature Structural and Molecular Biology*, **12**(11), 965–971.

Stark, G. R., & Taylor, W. R. (2004). Analyzing the G2/M checkpoint. *Methods in Molecular Biology*, **280**, 51–82.

Stark, M., Hiraga, S., & Donaldson, A. (2015). Protein phosphatases and DNA replication initiation. In D. Kaplan, ed., *Initiation of DNA Replication in Eukaryotes*, USA: Springer, pp. 461–477.

Stinchcomb, D. T., Struhl, K., & Davis, R. W. (1979). Isolation and characterisation of a yeast chromosomal replicator. *Nature*, **282**(5734), 39–43.

Stodola, J. L., & Burgers, P. M. (2016). Resolving individual steps of Okazaki-fragment maturation at a millisecond timescale. *Nature Structural and Molecular Biology*, **23**(5), 402–408.

Storici, F., Koh, K. D., Hesselberth, J., & Storici, F. (2015). Ribose-seq: ribonucleotides in DNA to Illumina library. *Protocol Exchange*. doi:10.1038/protex.2015.044

Stroud, H., Otero, S., Desvoyes, B., Ramírez-Parra, E., Jacobsen, S. E., & Gutierrez, C. (2012). Genome-wide analysis of histone H3.1 and H3.3 variants in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(14), 5370–5375.

Struhl, K., Stinchcomb, D. T., Scherer, S., & Davis, R. W. (1979). High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. *Proceedings of the National Academy of Sciences*, **76**(3), 1035–1039.

Stuckey, R., García-Rodríguez, N., Aguilera, A., & Wellinger, R. E. (2015). Role for RNA:DNA hybrids in origin-independent replication priming in a eukaryotic system. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(18), 5779–5784.

Sun, J., Kawakami, H., Zech, J., Speck, C., Stillman, B., & Li, H. (2012). Cdc6-induced conformational changes in ORC bound to origin DNA revealed by cryo-electron microscopy. *Structure*, **20**(3), 534–544.

Supek, F., & Lehner, B. (2015). Differential DNA mismatch repair underlies mutation rate variation across the human genome. *Nature*, **521**(7550), 81–84.

Suryadinata, R., Sadowski, M., & Sarcevic, B. (2010). Control of cell cycle progression by phosphorylation of cyclin-dependent kinase (CDK) substrates. *Bioscience Reports*, **30**(4), 243–255.

Tada, S., Li, A., Maiorano, D., Méchali, M., & Blow, J. J. (2001). Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. *Nature Cell Biology*, **3**(2), 107–113.

Taddei, A., & Gasser, S. M. (2004). Multiple pathways for telomere tethering: Functional implications of subnuclear position for heterochromatin formation. *Biochimica et Biophysica Acta - Gene Structure and Expression*, **1677**(1–3), 120–128.

Tak, Y. S., Tanaka, Y., Endo, S., Kamimura, Y., & Araki, H. (2006). A CDK-catalysed regulatory phosphorylation for formation of the DNA replication complex Sld2-Dpb11. *EMBO Journal*, **25**(9), 1987–1996.

Takahashi, S., Sakamoto, A., Sato, S., Kato, T., Tabata, S., & Tanaka, A. (2005). Roles of Arabidopsis AtREV1 and AtREV7 in translesion synthesis. *Plant Physiology*, **138**(2), 870–881.

Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A., & Araki, H. (2003). GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes and Development*, **17**(9), 1153–1165.

Tanaka, S., & Diffley, J. F. X. (2002). Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. *Nature Cell Biology*, **4**(3), 198–207.

Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y., & Araki, H. (2007). CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature*, **445**(7125), 328–332.

Tanaka, T., & Nasmyth, K. (1998). Association of RPA with chromosomal replication origins requires an Mcm protein, and is regulated by Rad53, and cyclin- and Dbf4-dependent kinases. *EMBO Journal*, **17**(17), 5182–5191.

Tardat, M., Brustel, J., Kirsh, O., Lefevbre, C., Callanan, M., Sardet, C., & Julien, E. (2010). The histone H4 Lys 20 methyltransferase PR-Set7 regulates replication

origins in mammalian cells. *Nature Cell Biology*, **12**(11), 1086–1093.

Tatsumi, Y., Sugimoto, N., Yugawa, T., Narisawa-Saito, M., Kiyono, T., & Fujita, M. (2006). Deregulation of Cdt1 induces chromosomal damage without rereplication and leads to chromosomal instability. *Journal of Cell Science*, **119**(15), 3128–3140.

Taylor, J. H. (1960). Asynchronous duplication of chromosomes in cultured cells of Chinese hamster. *The Journal of Biophysical and Biochemical Cytology*, **7**, 455–464.

Teixeira, M. T., & Gilson, E. (2005). Telomere maintenance, function and evolution: The yeast paradigm. *Chromosome Research*, **13**(5), 535–548.

Theis, J. F., Dershowitz, A., Irene, C., Maciariello, C., Tobin, M. L., Liberi, G., Tabrizifard, S., Korus, M., Fabiani, L., & Newlon, C. S. (2007). Identification of mutations that decrease the stability of a fragment of Saccharomyces cerevisiae chromosome III lacking efficient replicators. *Genetics*, **177**(3), 1445–1458.

Theis, J. F., Irene, C., Dershowitz, A., Brost, R. L., Tobin, M. L., di Sanzo, F. M., Wang, J. Y., Boone, C., & Newlon, C. S. (2010). The DNA damage response pathway contributes to the stability of chromosome III derivatives lacking efficient replicators. *PLoS Genetics*, **6**(12), 1–14.

Theis, J. F., & Newlon, C. S. (1997). The ARS309 chromosomal replicator of Saccharomyces cerevisiae depends on an exceptional ARS consensus sequence. *Proceedings of the National Academy of Sciences of the United States of America*, **94**(20), 10786–10791.

Theis, J. F., & Newlon, C. S. (2001). Two compound replication origins in Saccharomyces cerevisiae contain redundant origin recognition complex binding sites. *Molecular and Cellular Biology*, **21**(8), 2790–2801.

Ticau, S., Friedman, L. J., Ivica, N. A., Gelles, J., & Bell, S. P. (2015). Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. *Cell*, **161**(3), 513–525.

Treco, D. A., & Winston, F. (2008). Growth and manipulation of yeast. In *Current*

Protocols in Molecular Biology, pp. 1–12.

Tsakraklides, V., & Bell, S. P. (2010). Dynamics of pre-replicative complex assembly. *Journal of Biological Chemistry*, **285**(13), 9437–9443.

Tsankov, A., Yanagisawa, Y., Rhind, N., Regev, A., & Rando, O. J. (2011). Evolutionary divergence of intrinsic and trans-regulated nucleosome positioning sequences reveals plastic rules for chromatin organization. *Genome Research*, **21**(11), 1851–1862.

Tsurimoto, T., & Stillman, B. (1991). Replication factors required for SV40 DNA replication in vitro. II. Switching of DNA polymerase alpha and delta during initiation of leading and lagging strand synthesis. *Journal of Biological Chemistry*, **266**(3), 1961–1968.

Tsutakawa, S. E., Lafrance-Vanasse, J., & Tainer, J. A. (2014). The cutting edges in DNA repair, licensing, and fidelity: DNA and RNA repair nucleases sculpt DNA to measure twice, cut once. *DNA Repair*, **19**, 95–107.

Tye, B. K., Nyman, P. O., Lehman, I. R., Hochhauser, S., & Weiss, B. (1977). Transient accumulation of Okazaki fragments as a result of uracil incorporation into nascent DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **74**(1), 154–157.

Ullsperger, C. J., Vologodskii, A. V., & Cozzarelli, N. R. (1995). Unlinking of DNA by Topoisomerases During DNA Replication. In F. Eckstein & D. M. J. Lilley, eds., *Nucleic Acids and Molecular Biology*, Nucleic Ac, Berlin, Heidelberg: Springer, pp. 115–142.

Umek, R. M., & Kowalski, D. (1988). The ease of DNA unwinding as a determinant of initiation at yeast replication origins. *Cell*, **52**(4), 559–567.

Usongo, V., Martel, M., Balleydier, A., & Drolet, M. (2016). Mutations reducing replication from R-loops suppress the defects of growth, chromosome segregation and DNA supercoiling in cells lacking topoisomerase I and RNase HI activity. *DNA Repair*, **40**, 1–17.

Vafabakhsh, R., & Ha, T. (2012). Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cyclization. *Science*, **337**(6098), 1097–1101.

Van Houten, J. V, & Newlon, C. S. (1990). Mutational analysis of the consensus sequence of a replication origin from yeast chromosome III. *Molecular and Cellular Biology*, **10**(8), 3917–3925.

van Mourik, P. M., de Jong, J., Sharma, S., Kavšek, A., Chabes, A., & Chang, M. (2018). Upregulation of dNTP levels after telomerase inactivation influences telomerase-independent telomere maintenance pathway choice in Saccharomyces cerevisiae. *G3: Genes, Genomes, Genetics*, **8**(8), 2551–2558.

Vashee, S., Cvetic, C., Lu, W., Simancek, P., Kelly, T. J., & Walter, J. C. (2003). Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes and Development*, **17**(15), 1894–1908.

Vasianovich, Y., Harrington, L. A., & Makovets, S. (2014). Break-Induced replication requires DNA damage-induced phosphorylation of Pif1 and leads to telomere lengthening. *PLoS Genetics*, **10**(10), e1004679.

Vaziri, C., Saxena, S., Jeon, Y., Lee, C., Murata, K., Machida, Y., Wagle, N., Hwang, D. S., & Dutta, A. (2003). A p53-dependent checkpoint pathway prevents rereplication. *Molecular Cell*, **11**(4), 997–1008.

Vijayakumar, S., Chapados, B. R., Schmidt, K. H., Kolodner, R. D., Tainer, J. A., & Tomkinson, A. E. (2007). The C-terminal domain of yeast PCNA is required for physical and functional interactions with Cdc9 DNA ligase. *Nucleic Acids Research*, **35**(5), 1624–1637.

Villa, F., Simon, A. C., Ortiz Bazan, M. A., Kilkenny, M. L., Wirthensohn, D., Wightman, M., Matak-Vinkovíc, D., Pellegrini, L., & Labib, K. (2016). Ctf4 Is a hub in the eukaryotic replisome that links multiple CIP-box proteins to the CMG helicase. *Molecular Cell*, **63**(3), 385–396.

Virshup, D. M., & Shenolikar, S. (2009). From promiscuity to precision: protein phosphatases get a makeover. *Molecular Cell*, **33**(5), 537–545.

Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B. J., & Grunstein, M. (2002). Histone acetylation regulates the time of replication origin firing. *Molecular Cell*, **10**(5), 1223–1233.

Voichek, Y., Bar-Ziv, R., & Barkai, N. (2016). Expression homeostasis during DNA replication. *Science*, **351**(6277), 1087–1090.

von Meyenburg, K., Boye, E., Skarstad, K., Koppes, L., & Kogoma, T. (1987). Mode of initiation of constitutive stable DNA replication in RNase H-defective mutants of Escherichia coli K-12. *Journal of Bacteriology*, **169**(6), 2650–2658.

Voth, W. P., Olsen, A. E., Sbia, M., Freedman, K. H., & Stillman, D. J. (2005). ACE2, CBK1, and BUD4 in budding and cell separation. *Eukaryotic Cell*, **4**(6), 1018–1028.

Vujcic, M., Miller, C. A., & Kowalski, D. (1999). Activation of silent replication origins at autonomously replicating sequence elements near the HML locus in budding yeast. *Molecular and Cellular Biology*, **19**(9), 6098–6109.

Wallace, B. D., & Williams, R. S. (2014). Ribonucleotide triggered DNA damage and RNA-DNA damage responses. *RNA Biology*, **11**(11), 1340–1346.

Wallis, A. B. A., & Nieduszynski, C. A. (2018). Investigating the role of Rts1 in DNA replication initiation [version 1; referees: 2 approved]. *Wellcome Open Research*, **3**, 23.

Walsh, R. M., & Martin, P. A. (1977). Growth of Saccharomyces cerevisiae and Saccharomyces uvarum in a tempearture gradient incubator. *Journal of the Institute of Brewing*, **83**(3), 169–172.

Wang, G., Tong, X., Weng, S., & Zhou, H. (2012). Multiple phosphorylation of Rad9 by CDK is required for DNA damage checkpoint activation. *Cell Cycle*, **11**(20), 3792–3800.

Wanrooij, P. H., & Burgers, P. M. (2015). Yet another job for Dna2: Checkpoint activation. *DNA Repair*, **32**, 17–23.

Watson, J. D., & Crick, F. H. C. (1953). Molecular structure of nucleic acids. *Nature*, **171**(4356), 737–738.

Weinert, T. A., & Hartwell, L. H. (1988). The RAD9 gene controls the cell cycle response to DNA damage in saccharomyces cerevisiae. *Science*, **241**(4863), 317–322.

White, E. J., Emanuelsson, O., Scalzo, D., Royce, T., Kosak, S., Oakeley, E. J., Weissman, S., Gerstein, M., Groudine, M., Snyder, M., & Schübeler, D. (2004). DNA replication-timing analysis of human chromosome 22 at high resolution and different developmental states. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(51), 17771–17776.

Williams, G. H., Romanowski, P., Morris, L., Madine, M., Mills, A. D., Stoeber, K., Marr, J., Laskey, R. A., & Coleman, N. (1998). Improved cervical smear assessment using antibodies against proteins that regulate DNA replication. *Proceedings of the National Academy of Sciences of the United States of America*, **95**(25), 14932–14937.

Williams, J. S., Clausen, A. R., Lujan, S. A., Marjavaara, L., Clark, A. B., Burgers, P. M., Chabes, A., & Kunkel, T. A. (2015). Evidence that processing of ribonucleotides in DNA by topoisomerase 1 is leading-strand specific. *Nature Structural and Molecular Biology*. doi:10.1038/nsmb.2989

Wilson, M. A., Kwon, Y., Xu, Y., Chung, W. H., Chi, P., Niu, H., Mayle, R., Chen, X., Malkova, A., Sung, P., & Ira, G. (2013). Pif1 helicase and Polδ promote recombination-coupled DNA synthesis via bubble migration. *Nature*, **502**(7471), 393–396.

Wohlschlegel, J. A., Dwyer, B. T., Dhar, S. K., Cvetic, C., Walter, J. C., & Dutta, A. (2000). Inhibition of eukaryotic DNA replication by geminin binding to Cdt1. *Science*, **290**(5500), 2309–2312.

Wolfe, B. A., McDonald, W. H., Yates, J. R. R., & Gould, K. L. (2006). Phospho-Regulation of the Cdc14/Clp1 Phosphatase Delays Late Mitotic Events in S. pombe. *Developmental Cell*, **11**(3), 423–430.

Woodfine, K., Fiegler, H., Beare, D. M., Collins, J. E., McCann, O. T., Young, B. D., Debernardi, S., Mott, R., Dunham, I., & Carter, N. P. (2004). Replication timing of

the human genome. *Human Molecular Genetics*, **13**(2), 191–202.

Worning, P., Jensen, L. J., Hallin, P. F., Stærfeldt, H. H., & Ussery, D. W. (2006). Origin of replication in circular prokaryotic chromosomes. *Environmental Microbiology*, **8**(2), 353–361.

Wu, P., Takai, H., & De Lange, T. (2012a). Telomeric 3′ overhangs derive from resection by Exo1 and apollo and fill-in by POT1b-associated CST. *Cell*, **150**(1), 39–52.

Wu, Z., Liu, H., Liu, J., Liu, X., & Xiang, H. (2012b). Diversity and evolution of multiple *orc/cdc6*-adjacent replication origins in haloarchaea. *BMC Genomics*, **13**(1), 478.

Wu, Z., Liu, J., Yang, H., & Xiang, H. (2014). DNA replication origins in archaea. *Frontiers in Microbiology*, **5**, 179.

Xu, J., Yanagisawa, Y., Tsankov, A. M., Hart, C., Aoki, K., Kommajosyula, N., Steinmann, K. E., Bochicchio, J., Russ, C., Regev, A., Rando, O. J., Nusbaum, C., Niki, H., Milos, P., Weng, Z., & Rhind, N. (2012). Genome-wide identification and characterization of replication origins by deep sequencing. *Genome Biology*, **13**(4), R27.

Xu, W., Aparicio, J. G., Aparicio, O. M., & Tavaré, S. (2006). Genome-wide mapping of ORC and Mcm2p binding sites on tiling arrays and identification of essential ARS consensus sequences in S. cerevisiae. *BMC Genomics*, **7**, 276.

Xu, X., Rochette, P. J., Feyissa, E. A., Su, T. V., & Liu, Y. (2009). MCM10 mediates RECQ4 association with MCM2-7 helicase complex during DNA replication. *EMBO Journal*, **28**(19), 3005–3014.

Yabuki, N., Terashima, H., & Kitada, K. (2002). Mapping of early firing origins on a replication profile of budding yeast. *Genes to Cells*, **7**(8), 781–789.

Yamagata, K., Kato, J. I., Shimamot, A., Goto, M., Furuichi, Y., & Ikeda, H. (1998). Bloom's and Werner's syndrome genes suppress hyperrecombination in yeast sgs1 mutant: Implication for genomic instability in human diseases. *Proceedings of the National Academy of Sciences of the United States of America*, **95**(15), 8733–8738.

Yamashita, M., Hori, Y., Shinomiya, T., Obuse, C., Tsurimoto, T., Yoshikawa, H., &

Shirahige, K. (1997). The efficiency and timing of initiation of replication of multiple replicons of Saccharomyces cerevisiae chromosome VI. *Genes to Cells*, **2**(11), 655–665.

Yamazaki, S., Ishii, A., Kanoh, Y., Oda, M., Nishito, Y., & Masai, H. (2012). Rif1 regulates the replication timing domains on the human genome. *The EMBO Journal*, **31**(18), 3667–3677.

Yan, Z., Fedorov, S. A., Mumby, M. C., & Williams, R. S. (2000). PR48, a novel regulatory subunit of Protein Phosphatase 2A, interacts with Cdc6 and modulates DNA replication in human cells. *Molecular and Cellular Biology*, **20**(3), 1021–1029.

Yang, H., Wu, Z., Liu, J., Liu, X., Wang, L., Cai, S., & Xiang, H. (2015). Activation of a dormant replication origin is essential for Haloferax mediterranei lacking the primary origins. *Nature Communications*, **6**, 8321.

Yang, Y., Gu, Q., Zhang, Y., Sasaki, T., Crivello, J., O'Neill, R. J., Gilbert, D. M., & Ma, J. (2018). Continuous-trait probabilistic model for comparing multi-species functional genomic data. *Cell Systems*, **7**(2), 208–218.

Yano, K., Uesono, Y., Yoshida, S., Kikuchi, A., Kashiwazaki, J., Mabuchi, I., & Kikuchi, Y. (2013). Mih1/Cdc25 is negatively regulated by Pkc1 in Saccharomyces cerevisiae. *Genes to Cells*, **18**(6), 425–441.

Yeeles, J. T. P., Deegan, T. D., Janska, A., Early, A., & Diffley, J. F. X. (2015). Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature*, **519**(7544), 431–435.

Yeeles, J. T. P., Janska, A., Early, A., & Diffley, J. F. X. (2016). How the eukaryotic replisome achieves rapid and efficient DNA replication. *Molecular Cell*, **65**(1), 105–116.

Yuzhakov, A., Kelman, Z., Hurwitz, J., & O'Donnell, M. (1999). Multiple competition reactions for RPA order the assembly of the DNA polymerase δ holoenzyme. *EMBO Journal*, **18**(21), 6189–6199.

Zawilak-Pawlik, A., Kois, A., Majka, J., Jakimowicz, D., Smulczyk-Krawczyszyn, A.,

Messer, W., & Zakrzewska-Czerwińska, J. (2005). Architecture of bacterial replication initiation complexes: Orisomes from four unrelated bacteria. *Biochemical Journal*, **389**(2), 471–481.

Zechiedrich, E. L., Khodursky, A. B., & Cozzarelli, N. R. (1997). Topoisomerase IV, not gyrase, decatenates products of site-specific recombination in Escherichia coli. *Genes and Development*, **11**(19), 2580–2592.

Zegerman, P. (2015). Evolutionary conservation of the CDK targets in eukaryotic DNA replication initiation. *Chromosoma*, **124**(3), 309–321.

Zegerman, P., & Diffley, J. F. X. (2007). Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature*, **445**(7125), 281–285.

Zhang, J. M., Yadav, T., Ouyang, J., Lan, L., & Zou, L. (2019). Alternative lengthening of telomeres through two distinct break-induced replication pathways. *Cell Reports*, **26**(4), 955–968.

Zhu, W., Ukomadu, C., Jha, S., Senga, T., Dhar, S. K., Wohlschlegel, J. A., Nutt, L. K., Kornbluth, S., & Dutta, A. (2007). Mcm10 and And-1/CTF4 recruit DNA polymerase α to chromatin for initiation of DNA replication. *Genes and Development*, **21**(18), 2288–2299.

Zou, L., & Stillman, B. (2000). Assembly of a complex containing Cdc45p, replication protein A, and Mcm2p at replication origins controlled by S-phase cyclin-dependent kinases and Cdc7p-Dbf4p kinase. *Molecular and Cellular Biology*, **20**(9), 3086–3096.

8 Appendices

8.1 Pu-seq analysis script specific to this study

```
nano ~/post_mapping_single_strand_file_generation_two_polymerases.sh

#!/bin/bash
#load modules for programmes
module load SAMTOOLS
module load BEDTOOLS
module load R
module load BEDOPS

#Ask for file to input
echo Enter epsilon file name, which should be a \.bam file
read mapped_file_e
echo Enter delta file name, which should be a \.bam file
read mapped_file_d
echo Enter output file name
read output_file
output_file_e=${output_file}_e
output_file_d=${output_file}_d
echo Enter location of R app from home e.g. R_folder/Pu-seq.R
read pathTOfile


#Select forward and reverse reads to use, determine ribonucleotide number, and
offset base pairs

samtools view -h -q 30 -F 3848 -f 99 $mapped_file_e | grep -v XS:i: | samtools view
-b - | bedtools genomecov -5 -d -ibam stdin | awk 'BEGIN {OFS="\t"} {if ($3>0)
print $1,$2-2,$2-1,$3}' | awk '{for (i=2; i<=NF; i++) if ($i<0) next } {print}' >
${output_file_e}_f_offset.bg
```

```
samtools view -h -q 30 -F 3848 -f 83 $mapped_file_e | grep -v XS:i: | samtools view
-b - | bedtools genomecov -5 -d -ibam stdin | awk 'BEGIN {OFS="\t"} {if ($3>0)
print $1,$2,$2+1,$3}' | awk '{for (i=2; i<=NF; i++) if ($i<0) next } {print}' >
${output_file_e}_r_offset.bg


samtools view -h -q 30 -F 3848 -f 99 $mapped_file_d | grep -v XS:i: | samtools view
-b - | bedtools genomecov -5 -d -ibam stdin | awk 'BEGIN {OFS="\t"} {if ($3>0)
print $1,$2-2,$2-1,$3}' | awk '{for (i=2; i<=NF; i++) if ($i<0) next } {print}' >
${output_file_d}_f_offset.bg

samtools view -h -q 30 -F 3848 -f 83 $mapped_file_d | grep -v XS:i: | samtools view
-b - | bedtools genomecov -5 -d -ibam stdin | awk 'BEGIN {OFS="\t"} {if ($3>0)
print $1,$2,$2+1,$3}' | awk '{for (i=2; i<=NF; i++) if ($i<0) next } {print}' >
${output_file_d}_r_offset.bg


#epsilon
#Correct for the offset, and generate base position files

bedtools nuc -fi /data/nieduszynski/RESOURCES/REFERENCE_GENOMES/sacCer3.fa -bed
${output_file_e}_f_offset.bg -seq | awk 'BEGIN {OFS="\t"} {print $1,$2+1,$3+1,$4}'
> ${output_file_e}_f_correct.bg

bedtools nuc -fi /data/nieduszynski/RESOURCES/REFERENCE_GENOMES/sacCer3.fa -bed
${output_file_e}_f_offset.bg -seq | awk 'BEGIN {OFS="\t"} {print $1,$2+1,$3+1,$14}'
> ${output_file_e}_f_bases.bg

bedtools nuc -fi /data/nieduszynski/RESOURCES/REFERENCE_GENOMES/sacCer3.fa -bed
${output_file_e}_r_offset.bg -seq | awk 'BEGIN {OFS="\t"} {print $1,$2+1,$3+1,$4}'
> ${output_file_e}_r_correct.bg

bedtools nuc -fi /data/nieduszynski/RESOURCES/REFERENCE_GENOMES/sacCer3.fa -bed
${output_file_e}_r_offset.bg -seq | awk 'BEGIN {OFS="\t"} {print $1,$2+1,$3+1,$14}'
| awk 'BEGIN {OFS="\t"}{if ($4=="G"){$4="c"}}{if ($4=="C"){$4="g"}}{if
```

```
($4=="T"){$4="a"}}{if($4=="A"){$4="t"} print $1,$2,$3,$4}' >
${output_file_e}_r_bases.bg


#delta
#Correct for the offset, and generate base position files

bedtools nuc -fi /data/nieduszynski/RESOURCES/REFERENCE_GENOMES/sacCer3.fa -bed
${output_file_d}_f_offset.bg -seq | awk 'BEGIN {OFS="\t"} {print $1,$2+1,$3+1,$4}'
> ${output_file_d}_f_correct.bg

bedtools nuc -fi /data/nieduszynski/RESOURCES/REFERENCE_GENOMES/sacCer3.fa -bed
${output_file_d}_f_offset.bg -seq | awk 'BEGIN {OFS="\t"} {print $1,$2+1,$3+1,$14}'
> ${output_file_d}_f_bases.bg

bedtools nuc -fi /data/nieduszynski/RESOURCES/REFERENCE_GENOMES/sacCer3.fa -bed
${output_file_d}_r_offset.bg -seq | awk 'BEGIN {OFS="\t"} {print $1,$2+1,$3+1,$4}'
> ${output_file_d}_r_correct.bg

bedtools nuc -fi /data/nieduszynski/RESOURCES/REFERENCE_GENOMES/sacCer3.fa -bed
${output_file_d}_r_offset.bg -seq | awk 'BEGIN {OFS="\t"} {print $1,$2+1,$3+1,$14}'
| awk 'BEGIN {OFS="\t"}{if ($4=="G"){$4="c"}}{if ($4=="C"){$4="g"}}{if
($4=="T"){$4="a"}}{if($4=="A"){$4="t"} print $1,$2,$3,$4}' >
${output_file_d}_r_bases.bg



#Create output file for one polymerase

bedtools unionbedg -header -names fo_e fo_d f_bp_e f_bp_d re_d re_e r_bp_d r_bp_e -
i ${output_file_e}_f_correct.bg ${output_file_d}_f_correct.bg
${output_file_e}_f_bases.bg ${output_file_d}_f_bases.bg
${output_file_d}_r_correct.bg ${output_file_e}_r_correct.bg
${output_file_d}_r_bases.bg ${output_file_e}_r_bases.bg -/- | awk 'BEGIN {OFS="\t"}
{if (last&&$1!=last){sum=0; diff=0} last=$1; sum_e+=$4+$9; diff_e+=$4-$9;
sum_d+=$5+$8; diff_d+=$5-$8; if ($4>=1) for_base=$6; if ($4==0) for_base=$7; if
```

```
($8>=1) for_base=$10; if ($8==0) rev_base=$11; if (NR==1) print
$1,$2,$3,$4,$5,"f_bp",$8,$9,"r_bp","Cumu_e","Diff_e","Cumu_d","Diff_d"; if (NR>=2)
print $1,$2,$3,$4,$5,for_base,$8,$9,rev_base,sum_e,diff_e,sum_d,diff_d}' >
${output_file}_ed_complete.tab

echo Your summary file is ${output_file}_ed_complete.tab


#Create files for Carr lab Shiny App:
bedtools map -a
/data/nieduszynski/RESOURCES/REFERENCE_GENOMES/genomeWindows/sacCer3_300bp.bed -b
${output_file_d}_r_correct.bg -c 4 | awk 'BEGIN {OFS="\t"; print
"chro",",","pos",",","count"};{mid=($2+$3)/2}{if($4 == "."){$4=0}}{print
$1,",",mid,",",$4}' > ${output_file_d}_r_300win.csv

bedtools map -a
/data/nieduszynski/RESOURCES/REFERENCE_GENOMES/genomeWindows/sacCer3_300bp.bed -b
${output_file_e}_r_correct.bg -c 4 | awk 'BEGIN {OFS="\t"; print
"chro",",","pos",",","count"};{mid=($2+$3)/2}{if($4 == "."){$4=0}}{print
$1,",",mid,",",$4}' > ${output_file_e}_r_300win.csv

bedtools map -a
/data/nieduszynski/RESOURCES/REFERENCE_GENOMES/genomeWindows/sacCer3_300bp.bed -b
${output_file_d}_f_correct.bg -c 4 | awk 'BEGIN {OFS="\t"; print
"chro",",","pos",",","count"};{mid=($2+$3)/2}{if($4 == "."){$4=0}}{print
$1,",",mid,",",$4}' > ${output_file_d}_f_300win.csv

bedtools map -a
/data/nieduszynski/RESOURCES/REFERENCE_GENOMES/genomeWindows/sacCer3_300bp.bed -b
${output_file_e}_f_correct.bg -c 4 | awk 'BEGIN {OFS="\t"; print
"chro",",","pos",",","count"};{mid=($2+$3)/2}{if($4 == "."){$4=0}}{print
$1,",",mid,",",$4}' > ${output_file_e}_f_300win.csv
```

```
printf "\nCarr lab shiny app files
are:\n${output_file_d}_r_300win.csv\n${output_file_e}_r_300win.csv\n${output_file_d
}_f_300win.csv\n${output_file_e}_f_300win.csv\n\n"


#Remove excess files so space isn't taken up:

rm
{${output_file_e}_{f_offset.bg,f_bases.bg,r_offset.bg,r_bases.bg},${output_file_d}_
{f_offset.bg,f_bases.bg,r_offset.bg,r_bases.bg}}


#Change csv files to wig files for visualisation in IGV

Rscript ~/${pathTOfile} --vanilla --slave --args --prefix ${output_file} --df
${output_file_d}_f_300win.csv --dr ${output_file_d}_r_300win.csv --ef
${output_file_e}_f_300win.csv --er ${output_file_e}_r_300win.csv -ma1 3 ma2 3 -ori

printf "Wig files to use in IGV are:\n${output_file}_Pol-
epsilon.watson.wig\n${output_file}_Pol-delta.watson.wig\n${output_file}_Pol-
epsilon.crick.wig\n${output_file}_Pol-delta.crick.wig\n\n"


#Create bed files to use in R for visualisation using Phoebe's script

wig2bed --zero-indexed < ${output_file}_Pol-epsilon.watson.wig >
${output_file}_Pol-epsilon.watson.bed
wig2bed --zero-indexed < ${output_file}_Pol-delta.watson.wig > ${output_file}_Pol-
delta.watson.bed
wig2bed --zero-indexed < ${output_file}_Pol-epsilon.crick.wig > ${output_file}_Pol-
epsilon.crick.bed
wig2bed --zero-indexed < ${output_file}_Pol-delta.crick.wig > ${output_file}_Pol-
delta.crick.bed
```

```
printf "Bed files to use for R visualisation are:\n${output_file}_Pol-epsilon.watson.bed\n${output_file}_Pol-delta.watson.bed\n${output_file}_Pol-epsilon.crick.bed\n${output_file}_Pol-delta.crick.bed\n\n"
```

8.2 Wallis and Nieduszynski, 2018

A copy of the published work:

Wallis, ABA & Nieduszynski, CA (2018) Investigating the role of Rts1 in DNA replication initiation [version 1; referees: 2 approved], Wellcome Open Research 3 (23)

DOI: 10.12688/wellcomeopenres.13884.1

Downloaded **13th September 2019**

Check for updates

RESEARCH ARTICLE

Investigating the role of Rts1 in DNA replication initiation [version 1; peer review: 2 approved]

Ana B.A. Wallis, Conrad A. Nieduszynski

Sir William Dunn School of Pathology, University of Oxford, Oxford, Oxfordshire, OX1 3RE, UK

v1 **First published:** 06 Mar 2018, **3**:23 (https://doi.org/10.12688/wellcomeopenres.13884.1)

Latest published: 06 Mar 2018, **3**:23 (https://doi.org/10.12688/wellcomeopenres.13884.1)

Abstract

Background: Understanding DNA replication initiation is essential to understand the mis-regulation of replication seen in cancer and other human disorders. DNA replication initiates from DNA replication origins. In eukaryotes, replication is dependent on cell cycle kinases which function during S phase. Dbf4-dependent kinase (DDK) and cyclin-dependent kinase (CDK) act to phosphorylate the DNA helicase (composed of mini chromosome maintenance proteins: Mcm2-7) and firing factors to activate replication origins. It has recently been found that Rif1 can oppose DDK phosphorylation. Rif1 can recruit protein phosphatase 1 (PP1) to dephosphorylate MCM and restricts origin firing. In this study, we investigate a potential role for another phosphatase, protein phosphatase 2A (PP2A), in regulating DNA replication initiation. The PP2A regulatory subunit Rts1 was previously identified in a large-scale genomic screen to have a genetic interaction with *ORC2* (a DNA replication licensing factor). Deletion of *RTS1* synthetically rescued the temperature-sensitive (ts-) phenotype of *ORC2* mutants.

Methods: We deleted *RTS1* in multiple ts-replication factor *Saccharomyces cerevisiae* strains, including *ORC2*. Dilution series assays were carried out to compare qualitatively the growth of double mutant Δ*rts1* ts-replication factor strains relative to the respective single mutant strains.

Results: No synthetic rescue of temperature-sensitivity was observed. Instead we found an additive phenotype, indicating gene products function in separate biological processes. These findings are in agreement with a recent genomic screen which found that *RTS1* deletion in several ts-replication factor strains led to increased temperature-sensitivity.

Conclusions: We find no evidence that Rts1 is involved in the dephosphorylation of DNA replication initiation factors.

Keywords

DNA Replication, Phosphatases, Rts1, Rif1, Orc2

Open Peer Review

Reviewer Status ✓ ✓

	Invited Reviewers	
	1	**2**
version 1 published 06 Mar 2018	✓ report	✓ report

1 **Gilles Fischer**, Pierre and Marie Curie University (UPMC)-Sorbonne University, Paris, France

2 **David Shore**, University of Geneva, Geneva, Switzerland
Maksym Shyian, University of Geneva, Geneva, Switzerland

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Conrad A. Nieduszynski (conrad.nieduszynski@path.ox.ac.uk)

Author roles: Wallis ABA: Conceptualization, Investigation, Project Administration, Visualization, Writing – Original Draft Preparation; **Nieduszynski CA**: Conceptualization, Funding Acquisition, Project Administration, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: Wellcome Trust [110064] grant to CN. Biotechnology and Biological Sciences Research Council [BB/M011224/1] grant funded AW.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2018 Wallis ABA and Nieduszynski CA. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Wallis ABA and Nieduszynski CA. **Investigating the role of Rts1 in DNA replication initiation [version 1; peer review: 2 approved]** Wellcome Open Research 2018, **3**:23 (https://doi.org/10.12688/wellcomeopenres.13884.1)

First published: 06 Mar 2018, **3**:23 (https://doi.org/10.12688/wellcomeopenres.13884.1)

Introduction

Errors during DNA replication can lead to aneuploidy and DNA damage (Passerini *et al.*, 2016). An insufficient concentration of replication factors can also lead to genomic instability (Orr *et al.*, 2010). Therefore, it is important that cells ensure that a single round of DNA replication occurs in each cell cycle. DNA replication initiates from DNA replication origins (origins). In *Saccharomyces cerevisiae* origins are formed of an autonomously replicating sequence (ARS) which contains an 11bp ARS consensus sequence (ACS). Origins recruit the origin recognition complex (ORC) via the ACS, which in turn facilitates origin licensing. Origin licensing factors (Cdc6 and Cdt1) bind at the origin and allow the mini-chromosome maintenance (MCM) proteins to also bind. Post-licensing, firing factors (Cdc45, Sld2, Sld3, Dpb11) recruit the loading complex which contains GINS (a four-subunit complex), Cdc45 and the replicative polymerases (Polε, Polδ and Polα) (Yeeles *et al.*, 2015). Cdc45, MCM and GINS collectively form the CMG (Yeeles *et al.*, 2015). The CMG melts DNA, unwinding the DNA double helix to allow loading of the polymerases, to begin DNA replication. To prevent re-replication, origin licensing in eukaryotes is limited to G1 phase of the cell cycle, and origin firing is restricted to S phase (Blow & Dutta, 2005). In *S. cerevisiae*, loss of DNA re-replication control leads to genome instability including gene amplification (Green *et al.*, 2010).

The activities of licensing and firing factors are influenced by cell cycle kinases. For example, origin firing is dependent upon two kinases: the Dbf4-dependent kinase (DDK) and the cyclin-dependent kinase (CDK). DDK phosphorylates multiple chromatin-bound MCM subunits, including Mcm4 and Mcm6. Phosphorylation facilitates Sld3, Sld7 and Cdc45 binding. Subsequently, CDK phosphorylates Sld3 and Sld2, which then recruits the loading complex (Zegerman, 2015), which leads to origin firing.

However, the kinase-driven view of replication initiation outlined above is now known to be incomplete (Davé *et al.*, 2014). A role for dephosphorylation in controlling DNA replication initiation was established recently (Davé *et al.*, 2014; Hiraga *et al.*, 2014; Mattarocci *et al.*, 2014; Poh *et al.*, 2014). The Rap1-interacting factor (Rif1) is able to recruit protein phosphatase 1 (PP1) to MCM subunits and dephosphorylate them (Davé *et al.*, 2014; Hiraga *et al.*, 2014; Mattarocci *et al.*, 2014; Poh *et al.*, 2014). A greater DDK concentration is therefore required to promote origin firing, since the MCM phosphorylation rate must exceed its dephosphorylation rate. Conversely, DDK can bind directly to Rif1 and inhibit its interaction with PP1 (Hiraga *et al.*, 2014). Therefore, as DDK levels increase during S phase, MCM phosphorylation is promoted and dephosphorylation is inhibited. The resulting feedback loop allows for a rapid switch from low MCM phosphorylation in G1 to high MCM phosphorylation in S phase.

Rif1-PP1 involvement in DNA replication control appears to be conserved throughout eukaryotes, both *Xenopus* egg extract and HeLa cell studies support the findings in yeast (Poh *et al.*, 2014; Yamazaki *et al.*, 2012). Additionally, there is evidence that Rif1-PP1 controls further aspects of DNA replication initiation. For example, in yeast, Rif1-PP1 may antagonise CDK phosphorylation (Stark *et al.*, 2015). In *RIF1* null yeast strains phosphorylation of Sld3, but not Sld2, is increased (Mattarocci *et al.*, 2014). Deletion of *RIF1* can partially rescue the phenotype of temperature-sensitive (ts-) origin firing factor alleles including Dpb11, Cdc45 and Sld3 (Mattarocci *et al.*, 2014). In human cells, Rif1-PP1 is active during mitotic exit. Dephosphorylation of Orc2, an ORC subunit, allows the process of origin licensing to start again. Human Rif1-PP1 not only antagonises MCM phosphorylation, but also positively promotes DNA replication origin licensing (Hiraga *et al.*, 2017).

Before a role for Rif1-PP1 in DNA replication was described, Rif1 was known to be a telomere-associated protein, contributing to the late replication of telomeric regions (Lian *et al.*, 2011). Rif1 has also been implicated in a PP1-independent role in DNA replication at the whole genome level. The conserved replication timing of some genomic domains is altered in *RIF1* mutant cells due to disordered chromatin organisation. These observations led to a role for Rif1 in physically grouping similarly timed replication domains being described (Foti *et al.*, 2016).

The importance of Rif1-PP1 dephosphorylation raises the question of whether other phosphatases are implicated in DNA replication control. A large genomic screen for genetic interactions previously identified a potential synthetic rescue of mutant *ORC2* by additional *RTS1* deletion (Costanzo *et al.*, 2010). Rts1 is a regulatory subunit for the PP2A phosphatase, which has been previously implicated in DNA replication. PP2A antagonises the DNA damage checkpoint protein Chk1 (Petersen *et al.*, 2006), and its function is required for Cdc45 loading onto chromatin (Chou *et al.*, 2002; Peplowska *et al.*, 2014). Whether this interaction is direct, occurs via dephosphorylation of Sld3 (Guo *et al.*, 2015), or uses another protein complex (Chowdhury *et al.*, 2010) is as yet unclear. It has also been proposed that another regulatory subunit of PP2A, PR48, allows it to bind to and dephosphorylate the licensing factor Cdc6 during mitotic exit, promoting origin licensing during G1 (Yan *et al.*, 2000). Unlike PP1, which in humans is regulated by more than 90 different subunits, PP2A has only 13 regulatory subunits in humans, and 3 in yeast (Stark *et al.*, 2015).

Whilst Rif1 is associated with telomeres and late-replicating regions of DNA, Rts1 is associated with the protection of centromeres, which are known to replicate early (Mccarroll & Fangman, 1988). Rts1-PP2A is enriched at centromeres pre-anaphase promoting cell cycle progression after appropriate microtubule binding, and correct chromosome segregation (Peplowska *et al.*, 2014).

This study investigates a putative role for Rts1-PP2A, akin to Rif1-PP1, in controlling DNA replication licensing and firing. We use a panel of ts-replication factor mutants to screen for synthetic rescue by *RTS1* deletion. We find that *RTS1* deletion using classical genetics does not alleviate the lethality caused by inactivating origin initiation factors. Whilst the published synthetic rescue given by *rif1Δ* is confirmed in these strains, we

find an additive effect for *rts1Δ*. This indicates that two separate pathways are compromised. Although these data contradict the original genetic interactions screen (Costanzo *et al.*, 2010) they are in accordance with a more recent screen (Costanzo *et al.*, 2016), suggesting that *RTS1* deletion results in an enhanced (rather than alleviated) phenotype in some replication factor mutants.

Methods

Yeast strains and methods

Yeast strains were cultured both in liquid and on solid YPAD media (CCM1010 and CM0510 respectively; Formedium, Hunstanton, UK), and manipulated according to established practices (Treco & Winston, 2008). Most yeast strains used had a W303 background. However, strains from the *S. cerevisiae* genome deletion project (Giaever *et al.*, 2002) had an S288c background. All strains used are listed in Table 1.

In order to delete *S. cerevisiae* genes, the appropriate *KanMX* deletion cassettes from the SGDP were incorporated into a recipient strain by transformation. Deletion was confirmed by PCR spanning the deletion site. Oligonucleotide sequences are listed in Table 2. Ts-initiation factor mutant strains were confirmed by a lack of growth on solid YPAD plates at restrictive temperatures. Ts-initiation factor mutations with respective permissive and restrictive temperatures are listed in Table 3. Double mutant (ts-mutant / gene deletion) strains were confirmed by temperature-sensitivity and G418 resistance (400 µg/ml G418 disulfate salt; A1720-5G, Sigma-Aldrich, St Louis, MO, USA), relative to wild-type sister colonies.

Table 1. List of yeast strains. A list of yeast strains used in this study.

STRAIN	*GENOTYPE*	SOURCE
T7107	*MATa: RAD5, BUD4, leu2, ura3, trp1, ade2, his3*	T. Tanaka lab
45-1	*MATa: leu2-3, 112 ura3-52 ade2-1 lys2-801 cdc45-1*	C. Nieduszynski lab
CNY167	*MATa: ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 can1-100 Gal+ orc5-1*	C. Nieduszynski lab
AUY080	*MATa: ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 GAL+ ssd1,d2 RAD5 orc2-1*	C. Nieduszynski lab
K2539	*MATα: cdc9-1 Backcrossed three times to K699/K700*	T. Tanaka lab
dbf4-1	*MATa: ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 can1-100 ssd1-d2 Gal+ dbf4-1*	Tanaka & Nasmyth, 1998
YKB2	*MATa: leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3- 11,15, cdc7-4*	Mattarocci *et al.*, 2014
YYK32	*MATa: leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3- 11,15, cdc45-27, bar1Δ::hisG*	Mattarocci *et al.*, 2014
YYK14	*MATa: leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3- 11,15, sld3-4, bar1Δ::hisG*	Mattarocci *et al.*, 2014
YNIG63(2)	*MATa: leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3- 11,15 , dpb11-24, bar1Δ::hisG*	Mattarocci *et al.*, 2014
YCH175	*MATα: ho, ade2, trp1, can1, leu2, his3, GAL, psi + W303-1; cdc6-1*	Mattarocci *et al.*, 2014
YOR014W	*rts1Δ::KanMX S288c*	Giaever *et al.*, 2002
YBR275C	*rif1Δ::KanMX S288c*	Giaever *et al.*, 2002
YDR007W	*trp1Δ::KanMX S288c*	Giaever *et al.*, 2002
ACY001	W303 *MATα rts1Δ::kanMX*	This Study
ACY004	W303 *MATa rts1Δ::kanMX*	This Study
ACY007	W303 *MATα rif1Δ::kanMX*	This Study
ACY010	W303 *MATa rif1Δ::kanMX*	This Study
ACY013	W303 *MATα trp1Δ::kanMX*	This Study
ACY016	W303 *MATa trp1Δ::kanMX*	This Study
ACY036	W303 *Diploid orc2-1 rts1Δ::kanMX*	This Study
ACY113	W303 *MATα orc2-1 rts1Δ::kanMX*	This Study
ACY044	W303 *Diploid orc2-1 rif1Δ::kanMX*	This Study
ACY100	W303 *MATα orc2-1 rif1Δ::kanMX*	This Study

STRAIN	*GENOTYPE*	SOURCE
ACY079	W303 *Diploid cdc6-1 rts1Δ::kanMX*	This Study
ACY112	W303 *MATa cdc6-1 rts1Δ::kanMX*	This Study
ACY081	W303 *Diploid cdc6-1 rif1Δ::kanMX*	This Study
ACY148	W303 *MATa cdc6-1 rif1Δ::kanMX*	This Study
ACY035	W303 *Diploid cdc7-4 rts1Δ::kanMX*	This Study
ACY096	W303 *MATa cdc7-4 rts1Δ::kanMX*	This Study
ACY042	W303 *Diploid cdc7-4 rif1Δ::kanMX*	This Study
ACY139	W303 *MATa cdc7-4 rif1Δ::kanMX*	This Study
ACY071	W303 *Diploid dbf4-1 rts1Δ::kanMX*	This Study
ACY106	W303 *MATα dbf4-1 rts1Δ::kanMX*	This Study
ACY073	W303 *Diploid dbf4-1 rif1Δ::kanMX*	This Study
ACY104	W303 *MATα dbf4-1 rif1Δ::kanMX*	This Study
ACY031	W303 *Diploid cdc45-27 rts1Δ::kanMX*	This Study
ACY093	W303 *MATα cdc45-27 rts1Δ::kanMX*	This Study
ACY037	W303 *Diploid cdc45-27 rif1Δ::kanMX*	This Study
ACY142	W303 *MATa cdc45-27 rif1Δ::kanMX*	This Study
ACY019	W303 *Diploid cdc45-1 rts1Δ::kanMX*	This Study
ACY087	W303 *MATa cdc45-1 rts1Δ::kanMX*	This Study
ACY051	W303 *Diploid cdc45-1 rif1Δ::kanMX*	This Study
ACY145	W303 *MATa cdc45-1 rif1Δ::kanMX*	This Study
ACY025	W303 *Diploid cdc9-1 rts1Δ::kanMX*	This Study
ACY067	W303 *MATα cdc9-1 rts1Δ::kanMX*	This Study
ACY050	W303 *Diploid cdc9-1 rif1Δ::kanMX*	This Study
ACY120	W303 *MATα cdc9-1 rif1Δ::kanMX*	This Study
ACY069	W303 *Diploid dpb11-24 rts1Δ::kanMX*	This Study
ACY087	W303 *MATa dpb11-24 rts1Δ::kanMX*	This Study
ACY046	*W303 Diploid dpb11-24 rif1Δ::kanMX*	This Study
ACY123	W303 *MATα dpb11-24 rif1Δ::kanMX*	This Study

Dilution series assay

Strains were inoculated into YPAD liquid medium and cultured for 12–16 hours. Haploid cell concentration was inferred from attenuance measured at 600 nm (using a BioMate3 spectrophotometer; Thermofisher, Waltham, MA, USA). Cells were diluted initially to 10^7cells/ml, before serial 10-fold dilutions were prepared. Dilution spots of 5 µl were pipetted onto YPAD agar plates and incubated at stated temperatures. Control strains were included on each plate. After two days, plates were photographed and images were analysed qualitatively by observation of relative growth of yeast strains.

Results

RTS1 deletion does not synthetically rescue *orc2-1*

The combination of two mutations, which individually decrease cell fitness, can restore fitness if the genes have opposing effects (Synthetic Rescue, Figure 1). In the presence of a replication factor mutant, such as *orc2-1*, even at the permissive temperature origin firing can be reduced by as much as 30% (Shimada *et al.*, 2002), leading to growth deficiency. We first examined *RTS1* deletion in an *orc2-1* strain (Figure 2A), since a synthetic rescue phenotype has been reported (Costanzo *et al.*, 2010). A dilution series viability assay was used to assess synthetic rescue. We found that an *orc2-1 rts1Δ* strain had a more severe ts-phenotype that either the *orc2-1* or *rts1Δ* strains (Figure 2A). This additive effect indicates that the two genes are not acting within the same pathway. Conversely, a small synthetic rescue was observed in the *orc2-1 rif1Δ* strain (Figure 2A). It has previously been shown that *RIF1* deletion leads to slight synthetic rescue in *orc5-1* strains, consistent with this result (Mattarocci *et al.*, 2014).

The origin licensing factor Cdc6 is not opposed by *RTS1*

Since Rif1 recruits PP1 to oppose DDK phosphorylation, *RIF1* deletion provides limited or no rescue to the temperature sensitivity of pre-Replication Complex (pre-RC) factors mutants, which

Table 2. List of oligonucleotides. A list of oligonucleotides used in this study.

PRIMER 1	PRIMER 2	PRODUCT
AC0003 TTTTCAGTTCTTTGTGTTTTTCCTC	AC0004 TGATCCTTTAGAATGGAGAAGATTG	*rif1Δ::kanMX*
AC0005 TAAACCATCGTCGCCGTAA	AC0006 GGAAGAAGGAAAGCGAAAAGA	*rts1Δ::kanMX*
CA1118 CCATTACGCTCGTCATCAAA	AC0010 AAGAAACAAGAAGTCAACAGAAGG	Confirms 5' insertion of *rif1Δ::kanMX*
CA1117 GATAATGTCGGGCAATCAGG	AC0009 GCGGTAGCATTTCCATCATAA	Confirms 3' insertion of *rif1Δ::kanMX*
CA1118 CCATTACGCTCGTCATCAAA	AC0011 GGCATGTCAATACGTCTCGTT	Confirms 5' insertion of *rts1Δ::kanMX*
CA1117 GATAATGTCGGGCAATCAGG	AC0012 GGCAAGGTTTACGGAAAAGA	Confirms 3' insertion of *rts1Δ::kanMX*

Table 3. Temperature-sensitive mutation strains used in this study. Mutant forms of replication initiation factors, temperatures at which we observed phenotypes, and the study that originally reported each strain.

TS-REPLICATION FACTOR MUTATION	TEMPERATURE AT WHICH PHENOTYPE OBSERVED	TEMPERATURE AT WHICH NO PHENOTYPE OBSERVED	REFERENCE
orc2-1	30	23	Foss *et al.*, 1993
cdc6-1	33	30	Hartwell *et al.*, 1973
cdc7-4	30	23	Hartwell *et al.*, 1973
dbf4-1	32	30	Johnston & Thomas, 1982
cdc45-27	32	30	Kamimura *et al.*, 2001
cdc45-1	23	30	Moir *et al.*, 1982
cdc9-1	30	23	Hartwell *et al.*, 1973
dpb11-24	37	32	Tanaka *et al.*, 2007

function prior to DDK (Mattarocci *et al.*, 2014). Given that this study aimed to investigate a role for Rts1 in opposing the action of Orc2, we next looked for a genetic interaction between *RTS1* and another origin licensing factor: *CDC6*, which loads MCM. Deleting *RTS1* in the context of *cdc6*-1 showed no rescue relative to the original ts-strain (Figure 2B). Similarly, deletion of *RIF1* gave no synthetic rescue, consistent with published data (Mattarocci *et al.*, 2014). It has previously been shown that *rts1Δ* yeast strains are ts at 37°C (Auesukaree *et al.*, 2009; Shu & Hallberg, 1995), while *rif1Δ* strains are not (Mattarocci *et al.*, 2014). Our study confirmed both these phenotypes (Figure 2B).

Rts1 does not antagonise DDK phosphorylation

A potential role for Rts1-PP2A phosphatase in DNA replication initiation could be to oppose the action of a kinase. The established role for Rif1-PP1 in opposing DDK activity indicates that regulation of phosphorylation is key during this step of replication initiation. Therefore, *RTS1* was deleted in combination with ts-forms of both subunits of DDK (Cdc7 and Dbf4). However, *rts1Δ* had a slightly additive effect on temperature-sensitivity in both *cdc7-4* (Figure 3A) and *dbf4-1* (Figure 3B) strains. This was in stark contrast to the strong restoration of cell growth at elevated temperatures in *cdc7-4 rif1Δ* (Figure 3A) and *dbf4-1 rif1Δ* (Figure 3B) strains. This suggests that Rif1 and Rts1 do not have analogous roles in control of DNA replication initiation, and that Rts1-PP2A does not antagonise DDK activity.

Replication firing factors are not opposed by Rts1

MCM phosphorylation by DDK recruits the firing factor, Cdc45. The sequential recruitment of further firing factors (e.g. Dpb11) relies on phosphorylation by CDK. Therefore, Rts1-PP2A activity could be important following DDK activity, to limit CDK-induced origin firing. In order to test this hypothesis, *RTS1* was deleted in the context of ts- *cdc45-1* and *dpb11-24* firing factors. Unlike *rif1Δ*,

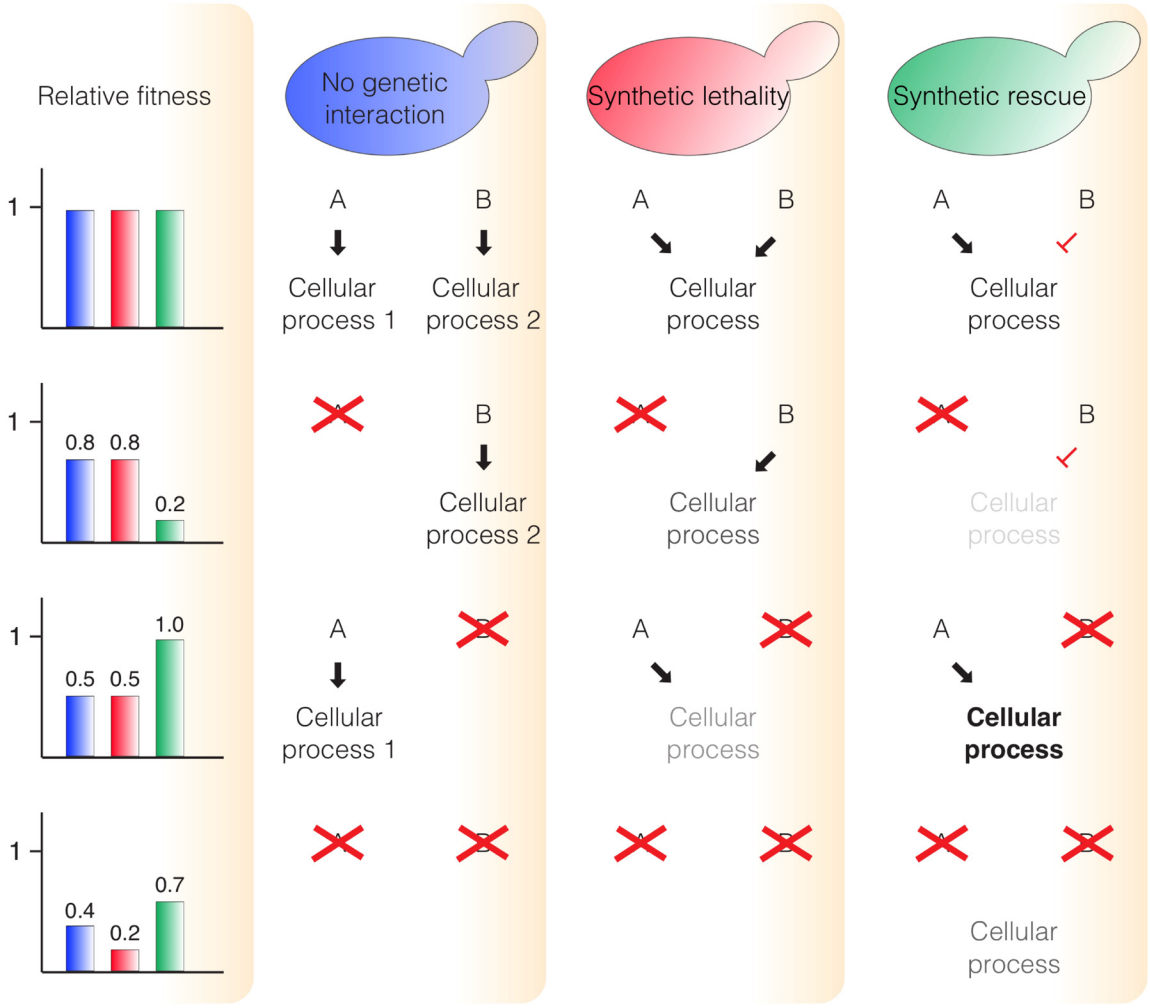


Figure 1. A summary of genetic interactions. Two genes, *A* and *B*, show genetic interactions as a result of the interacting functions of their products: A and B. When A and B function in different cellular processes, the relative fitness of an $A^{-}B^{-}$ double mutant is a product of the relative fitness of the two single mutants (no genetic interaction). If the double mutant strain has a lower than expected viability, it is described as synthetic lethality, indicating redundant functions for the two gene products in one cellular process. In contrast, a greater than expected viability (synthetic rescue) indicates that the gene products have opposing roles in a cellular process.

rts1Δ did not rescue *cdc45-27* or *dpb11-24* temperature-sensitivity (Figure 4AI, 4B). Conversely, *rts1Δ* led to an increased lethality in these strains, suggesting either an additive or synthetic lethality effect (Figure 1).

However, the cold-sensitive *cdc45-1* strain (restricted at 15°C) had no phenotypic rescue by either *rif1Δ* or *rts1Δ* (Figure 4AII). Instead, an additive effect was observed for both gene deletions, which was stronger in the case of *RTS1*. This additive effect of *rif1Δ* in cold-sensitive *cdc45-1* contradicts the known rescue of ts-*cdc45-27* by *rif1Δ* (Mattarocci *et al.*, 2014), a finding repeated in this study. Therefore, the significance of an additive effect of *rts1Δ* with *cdc45-1* is unclear.

Synthetic lethality of *RTS1* and *CDC9* ligase

Since the detrimental effect of *RTS1* deletion alongside ts-DNA replication initiation factors appeared to be most severe in factors which function later in the firing process, we hypothesised that Rts1 could play a role in DNA replication post-firing. Therefore, *RTS1* was deleted alongside a ts-*CDC9* ligase allele (*cdc9-1*). Cdc9 ligates lagging strand Okazaki fragments, aiding in DNA replication during elongation. *RTS1* deletion in a *cdc9-1* strain led to the greatest synthetic lethality, relative to *RTS1* deletion in the other ts-strains examined. No effect of *rif1Δ* was seen in *cdc9-1* cells. This was anticipated, since Rif1 is known to function during replication initiation. These data show that, if Rts1 plays a role in DNA replication, it is not akin to that

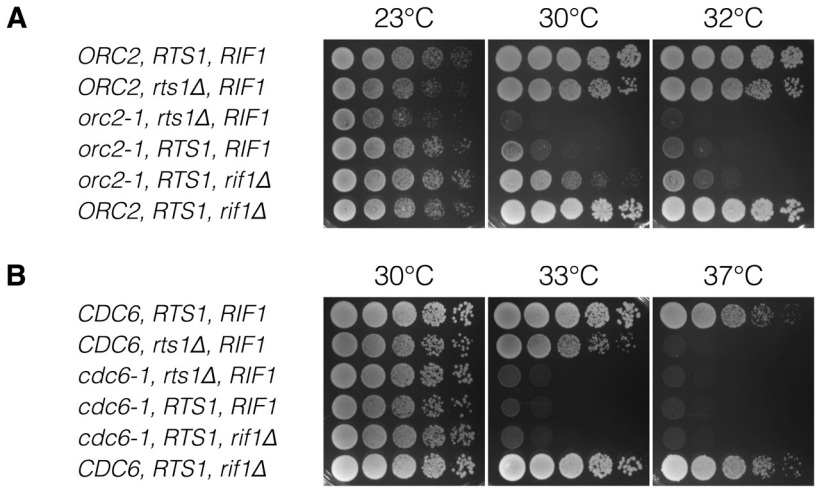


Figure 2. *RTS1* deletion does not suppress temperature-sensitivity of DNA replication origin licensing factor mutants. Budding yeast strains with ts-mutants of replication factors, together with either wild type, *rts1Δ* or *rif1Δ* were characterised by dilution viability assays. Wild type strains, without ts-replication factors, are shown at the top of each panel, as a control. (**A**) ORC subunit (*orc2-1*) is assayed. (**B**) A second pre-RC component, *cdc6-1* is assayed.

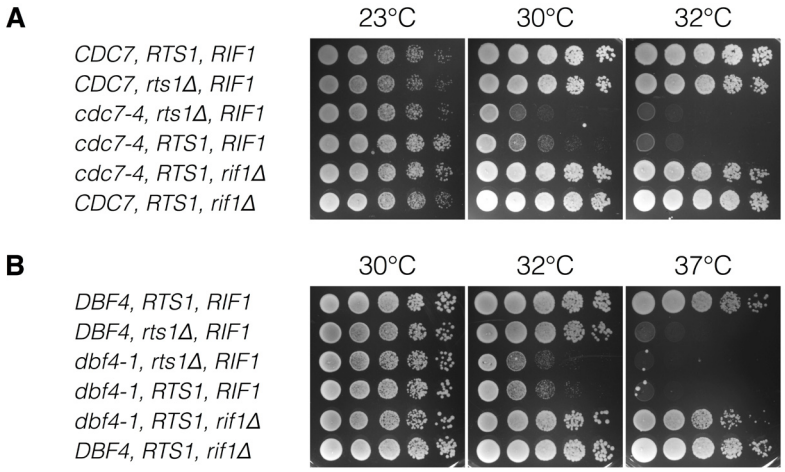


Figure 3. Unlike *rif1Δ*, *rts1Δ* cannot synthetically rescue ts-forms of DDK subunits. Budding yeast strains with ts-mutants of replication factors, together with either wild type, *rts1Δ* or *rif1Δ* were characterised by dilution viability assays. Wild type strains, without ts-replication factors, are shown at the top of each panel, as a control. Ts-forms of DDK subunits, Cdc7 (**A**) and Dbf4 (**B**) are assayed.

played by Rif1, and does not appear to oppose critical events leading to origin licensing or origin firing.

Discussion

In this study, *RTS1*, which encodes a regulatory subunit for the PP2A phosphatase, was deleted in the context of a range of ts-replication factor mutants. At no stage of DNA replication initiation (licensing, DDK-phosphorylation, and origin firing) did deletion of *RTS1* lead to synthetic rescue of ts-phenotypes. In contrast, deletion of *RIF1* was able to rescue ts-mutants of replication firing factors, and some replication licensing factors. Rif1 recruits PP1 phosphatase to DNA replication origins where it counteracts MCM phosphorylation by DDK. When replication origin firing is limited by mutant replication factors, removing this negative regulation allows more replication origins to fire, giving synthetic rescue. In contrast, *rts1Δ* leads to increased temperature-sensitivity when combined with *orc2-1*, *cdc7-4*, *cdc45-1* and *dpb11-24* mutant replication factors. The lack of synthetic rescue given by *rts1Δ* in these strains indicates that there is no evidence for a role for Rts1 in limiting origin firing, analogous to that played by Rif1.

Without a genetic interaction, combining mutations in genes involved in separate pathways will give an additive effect.

However, alone, *rts1Δ* has no growth inhibition at temperatures below 34°C. Therefore, the extent of reduced cell viability seen in some double mutant strains, such as *dpb11-24 rts1Δ* at 32°C (Figure 4), suggests synthetic lethality. This may be the result of non-specific protein instability after heat stress, in *rts1Δ* strains. Over-expression of *RTS1* can partially rescue lethality of a ts-*HPS60* allele (Shu & Hallberg, 1995). Hps60 is a mitochondrial protein that aids in protein refolding after heat stress (Shu & Hallberg, 1995). A reduced capability to maintain protein structure in heat stress conditions could explain the increased temperature sensitivity of unstable replication factor forms in *rts1Δ* cells. This could be analogous to the partial rescue of the ts-phenotype of *orc2-1* by mutations in the ubiquitin ligase *UBA1* (Shimada *et al.*, 2002). Therefore, investigation of heat stress in *rts1Δ* strains would be needed to elucidate the molecular mechanism.

The extent of the additional lethality given by *rts1Δ* alongside mutant ts-replication factors is inconsistent. If added lethality of *rts1Δ* depends on the function of the ts-factor, this could indicate a functional genetic interaction between that replication factor and *RTS1*. In DDK and pre-RC factor mutants (*cdc7-4, dbf4-1, cdc6-1*) there is either mildly increased temperature-sensitivity or no effect given by *rts1Δ* (Figure 2B, Figure 3). However, in post-DDK acting firing factors *cdc45-1* and *dpb11-24*, the observed increase in the ts-phenotype is larger (Figure 4). We cannot exclude the possibility that this effect is due to greater heat-instability of the *cdc45-1* and *dpb11-24* mutant replication factors. However, these

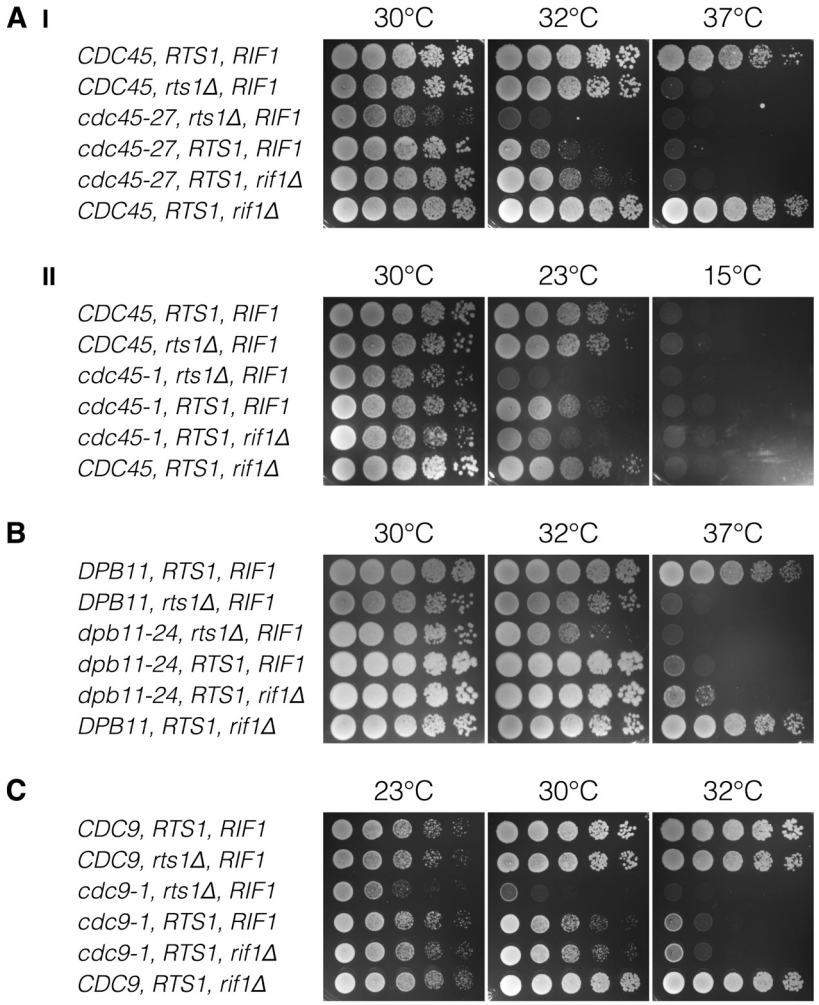


Figure 4. *RTS1* causes synthetic lethality with DNA replication firing factors and a DNA replication progression factor. Budding yeast strains with ts-mutants of replication factors, together with either wild type, *rts1Δ* or *rif1Δ* were characterised by dilution viability assays. Wild type strains, without ts-replication factors, are shown at the top of each panel, as a control. (**A**) Ts- (I) and cold- sensitive (II) forms of Cdc45, (**B**) the replication firing factor Dbp11, which functions post-DDK, and (**C**) Cdc9 ligase, are assayed.

data may suggest a role for Rts1 late in DNA replication initiation, demonstrating a genuine negative genetic interaction between *RTS1* and post-DDK replication firing factors.

Interestingly, the greatest synthetic lethality is seen between *RTS1* and a replication elongation factor: *CDC9*. Cdc9 is important for DNA replication progression and elongation rather than initiation. Accordingly, we find no synthetic rescue by *rif1Δ* in the *cdc9-1* strain (Figure 4). The observed synthetic lethality of *RTS1* deletion in a *cdc9-1* strain may provide evidence for a complementary role for Rts1 function in allowing replication fork progression. In *cdc9-1* strains, increased DNA damage is seen due to the collapse of replication forks. This results in double strand breaks (DSBs), and the DNA damage response (DDR) being activated. One of the ways to repair DSBs is via break-induced replication (BIR), which is activated in *cdc9-1* cells (Vasianovich *et al.*, 2014). A role for Rts1 in recruiting PP2A phosphatase to control phosphorylation steps in the DDR, potentially in BIR, could be hypothesised. In this instance, *rts1Δ* in a *cdc9-1* background would give increased lethality, since there would also be an impaired capacity for cells to repair *cdc9-1* dependent DSBs.

Evidence in support of the synthetic lethality of *RTS1* and DNA replication factors can be found on BioGRID, a summary of published genetic interactions in budding yeast. A combination of high and low throughput genetic interaction screens show that *RTS1* exhibits negative genetic interactions (a term reserved for genetic screens which show a more lethal phenotype in strains where two mutations are combined than in the respective single mutant strains) with *DBF4*, *CDC6* and *ORC6* and *DPB11* (Archambault *et al.*, 2005; Collins *et al.*, 2007; Costanzo *et al.*, 2016). However, we show here, for the first time, negative genetic interactions of *RTS1* with *CDC45* and *CDC9*.

Conclusions

Given the wealth of recent literature outlining the importance of Rif1 in opposing the actions of DDK kinase, it is clear that phosphatases play an important role in controlling DNA replication origin firing. However, we do not find evidence for an analogous role for PP2A, specifically via its regulatory subunit Rts1. Deletion of *RTS1* in combination with mutations in origin licensing factor genes, *ORC* and *CDC6*, showed little or no genetic interaction, providing no genetic evidence for Rts1-PP2A controlling DNA replication origin licensing. Further, we found no role for Rts1 in opposing DDK phosphorylation. However, we observed some level of increased temperature-sensitivity phenotype when *RTS1* was deleted in many of the replication initiation factor mutant strains, alluding to a potential synthetic lethality phenotype. Increased temperature-sensitivity was most pronounced in late-acting DNA replication firing factors Cdc45 and Dpb11. Additionally, an increased requirement for *RTS1* in a *cdc9-1* background was found. We speculate that a functional overlap between Rts1-PP2A and Cdc9 may exist via replication fork progression mechanisms. Rts1 may recruit PP2A during BIR, or the DDR, in response to DSBs. Over-expression of *RTS1* could potentially compensate for the increased replication fork collapse seen in *cdc9-1* mutants, giving synthetic rescue of temperature-sensitivity. Further studies would be needed to confirm this hypothesis.

Data availability

All data underlying the results are available as part of the article and no additional source data are required.

Competing interests

No competing interests were disclosed.

Grant information

Wellcome Trust [110064] grant to CN.

Biotechnology and Biological Sciences Research Council [BB/M011224/1] grant funded AW.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

We are grateful to Tomoyuki Tanaka and David Shore for sending us yeast strains, detailed in Table 1. We are also grateful to Carolin A. Müller, Dzmitry G. Batrakou, Rosemary H.C. Wilson and Emma D. Heron who gave feedback on the manuscript.

References

Archambault V, Ikui AE, Drapkin BJ, *et al.*: **Disruption of mechanisms that prevent rereplication triggers a DNA damage response.** *Mol Cell Biol.* 2005; **25**(15): 6707–6721.
PubMed Abstract | Publisher Full Text | Free Full Text

Auesukaree C, Damnernsawad A, Kruatrachue M, *et al.*: **Genome-wide identification of genes involved in tolerance to various environmental stresses in *Saccharomyces cerevisiae*.** *J Appl Genet.* 2009; **50**(3): 301–310.
PubMed Abstract | Publisher Full Text

Blow JJ, Dutta A: **Preventing re-replication of chromosomal DNA.** *Nat Rev Mol Cell Biol.* 2005; **6**(6): 476–486.
PubMed Abstract | Publisher Full Text | Free Full Text

Chou DM, Petersen P, Walter JC, *et al.*: **Protein phosphatase 2A regulates binding of Cdc45 to the prereplication complex.** *J Biol Chem.* 2002; **277**(43): 40520–40527.
PubMed Abstract | Publisher Full Text

Chowdhury A, Liu G, Kemp M, *et al.*: **The DNA unwinding element binding protein DUE-B interacts with Cdc45 in preinitiation complex formation.** *Mol Cell Biol.* 2010; **30**(6): 1495–1507.
PubMed Abstract | Publisher Full Text | Free Full Text

Collins SR, Miller KM, Maas NL, *et al.*: **Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map.** *Nature.* 2007; **446**(7137): 806–10.
PubMed Abstract | Publisher Full Text

Costanzo M, Baryshnikova A, Bellay J, *et al.*: **The genetic landscape of a cell.**

Science. 2010; **327**(5964): 425–31.
PubMed Abstract | Publisher Full Text | Free Full Text

Costanzo M, VanderSluis B, Koch EN, *et al.*: **A global genetic interaction network maps a wiring diagram of cellular function.** *Science.* 2016; **353**(6306): pii: aaf1420.
PubMed Abstract | Publisher Full Text | Free Full Text

Davé A, Cooley C, Garg M, *et al.*: **Protein Phosphatase 1 recruitment by Rif1 regulates DNA replication origin firing by counteracting DDK activity.** *Cell Rep.* 2014; **7**(1): 53–61.
PubMed Abstract | Publisher Full Text | Free Full Text

Foss M, McNally FJ, Laurenson P, *et al.*: **Origin recognition complex (ORC) in transcriptional silencing and DNA replication in S. cerevisiae.** *Science.* 1993; **262**(5141): 1838–44.
PubMed Abstract | Publisher Full Text

Foti R, Gnan S, Cornacchia D, *et al.*: **Nuclear Architecture Organized by Rif1 Underpins the Replication-Timing Program.** *Mol Cell.* 2016; **61**(2): 260–273.
PubMed Abstract | Publisher Full Text | Free Full Text

Giaever G, Chu AM, Ni L, *et al.*: **Functional profiling of the *Saccharomyces cerevisiae* genome.** *Nature.* 2002; **418**(6896): 387–391.
PubMed Abstract | Publisher Full Text

Green BM, Finn KJ, Li JJ: **Loss of DNA replication control is a potent inducer of gene amplification.** *Science.* 2010; **329**(5994): 943–6.
PubMed Abstract | Publisher Full Text | Free Full Text

Guo C, Kumagai A, Schlacher K, *et al.*: **Interaction of Chk1 with Treslin negatively regulates the initiation of chromosomal DNA replication.** *Mol Cell.* 2015; **57**(3): 492–505.
PubMed Abstract | Publisher Full Text | Free Full Text

Hartwell LH, Mortimer RK, Culotti J, *et al.*: **Genetic Control of the Cell Division Cycle in Yeast: V. Genetic Analysis of *cdc* Mutants.** *Genetics.* 1973; **74**(2): 267–286.
PubMed Abstract | Free Full Text

Hiraga S, Alvino GM, Chang F, *et al.*: **Rif1 controls DNA replication by directing Protein Phosphatase 1 to reverse Cdc7-mediated phosphorylation of the MCM complex.** *Genes Dev.* 2014; **28**(4): 372–383.
PubMed Abstract | Publisher Full Text | Free Full Text

Hiraga S, Ly T, Garzón J, *et al.*: **Human RIF1 and protein phosphatase 1 stimulate DNA replication origin licensing but suppress origin activation.** *EMBO Rep.* 2017; **18**(3) 403-419.
PubMed Abstract | Publisher Full Text

Johnston LH, Thomas AP: **The isolation of new DNA synthesis mutants in the yeast *Saccharomyces cerevisiae*.** *Mol Gen Genet.* 1982; **186**(3): 439–444.
PubMed Abstract | Publisher Full Text

Kamimura Y, Tak YS, Sugino A, *et al.*: **Sld3, which interacts with Cdc45 (Sld4), functions for chromosomal DNA replication in *Saccharomyces cerevisiae*.** *EMBO J.* 2001; **20**(8): 2097–2107.
PubMed Abstract | Publisher Full Text | Free Full Text

Lian HY, Robertson ED, Hiraga S, *et al.*: **The effect of Ku on telomere replication time is mediated by telomere length but is independent of histone tail acetylation.** *Mol Biol Cell.* 2011; **22**(10): 1753–1765.
PubMed Abstract | Publisher Full Text | Free Full Text

Mattarocci S, Shyian M, Lemmens L, *et al.*: **Rif1 controls DNA replication timing in yeast through the PP1 Phosphatase Glc7.** *Cell Rep.* 2014; **7**(1): 62–69.
PubMed Abstract | Publisher Full Text

Mccarroll RM, Fangman WL: **Time of replication of yeast centromeres and telomeres.** *Cell.* 1988; **54**(4): 505–513.
PubMed Abstract | Publisher Full Text

Moir D, Stewart SE, Osmond BC, *et al.*: **Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties, and pseudoreversion studies.** *Genetics.* 1982; **100**(4): 547–563.
PubMed Abstract | Free Full Text

Orr SJ, Gaymes T, Ladon D, *et al.*: **Reducing MCM levels in human primary T cells during the G_0-->G_1 transition causes genomic instability during the first cell cycle.** *Oncogene.* 2010; **29**(26): 3803–14.
PubMed Abstract | Publisher Full Text | Free Full Text

Passerini V, Ozeri-Galai E, de Pagter MS, *et al.*: **The presence of extra chromosomes leads to genomic instability.** *Nat Commun.* 2016; **7**: 10754.
PubMed Abstract | Publisher Full Text | Free Full Text

Peplowska K, Wallek AU, Storchova Z: **Sgo1 regulates both condensin and Ipl1/Aurora B to promote chromosome biorientation.** *PLoS Genet.* 2014; **10**(6): e1004411.
PubMed Abstract | Publisher Full Text | Free Full Text

Petersen P, Chou DM, You Z, *et al.*: **Protein phosphatase 2A antagonizes ATM and ATR in a Cdk2- and Cdc7-independent DNA damage checkpoint.** *Mol Cell Biol.* 2006; **26**(5): 1997–2011.
PubMed Abstract | Publisher Full Text | Free Full Text

Poh WT, Chadha GS, Gillespie PJ, *et al.*: ***Xenopus* Cdc7 executes its essential function early in S phase and is counteracted by checkpoint-regulated protein phosphatase 1.** *Open Biol.* 2014; **4**(1): 130138.
PubMed Abstract | Publisher Full Text | Free Full Text

Shimada K, Pasero P, Gasser SM: **ORC and the intra-S-phase checkpoint: A threshold regulates Rad53p activation in S phase.** *Genes Dev.* 2002; **16**(24): 3236–3252.
PubMed Abstract | Publisher Full Text | Free Full Text

Shu Y, Hallberg RL: **SCS1, a multicopy suppressor of hsp60-ts mutant alleles, does not encode a mitochondrially targeted protein.** *Mol Cell Biol.* 1995; **15**(10): 5618–5626.
PubMed Abstract | Publisher Full Text | Free Full Text

Stark M, Hiraga S, Donaldson A: **Protein phosphatases and DNA replication initiation.** In Kaplan D, ed. *Initiation of DNA Replication in Eukaryotes.* USA: Springer, 2015; 461–477.
Publisher Full Text

Tanaka S, Umemori T, Hirai K, *et al.*: **CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast.** *Nature.* 2007; **445**(7125): 328–332.
PubMed Abstract | Publisher Full Text

Tanaka T, Nasmyth K: **Association of RPA with chromosomal replication origins requires an Mcm protein, and is regulated by Rad53, and cyclin- and Dbf4-dependent kinases.** *EMBO J.* 1998; **17**(17): 5182–5191.
PubMed Abstract | Publisher Full Text | Free Full Text

Treco DA, Winston F: **Growth and manipulation of yeast.** *Curr Protoc Mol Biol.* 2008; **Chapter 13**(SUPPL. 82): Unit13.2.
PubMed Abstract | Publisher Full Text

Vasianovich Y, Harrington LA, Makovets S: **Break-induced replication requires DNA damage-induced phosphorylation of Pif1 and leads to telomere lengthening.** *PLoS Genet.* 2014; **10**(10): e1004679.
PubMed Abstract | Publisher Full Text | Free Full Text

Yamazaki S, Ishii A, Kanoh Y, *et al.*: **Rif1 regulates the replication timing domains on the human genome.** *EMBO J.* 2012; **31**(18): 3667–77.
PubMed Abstract | Publisher Full Text | Free Full Text

Yan Z, Fedorov SA, Mumby MC, *et al.*: **PR48, a novel regulatory subunit of protein phosphatase 2A, interacts with Cdc6 and modulates DNA replication in human cells.** *Mol Cell Biol.* 2000; **20**(3): 1021–1029.
PubMed Abstract | Publisher Full Text | Free Full Text

Yeeles JT, Deegan TD, Janska A, *et al.*: **Regulated eukaryotic DNA replication origin firing with purified proteins.** *Nature.* 2015; **519**(7544): 431–435.
PubMed Abstract | Publisher Full Text | Free Full Text

Zegerman P: **Evolutionary conservation of the CDK targets in eukaryotic DNA replication initiation.** *Chromosoma.* 2015; **124**(3): 309–321.
PubMed Abstract | Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 1

Reviewer Report 10 April 2018

https://doi.org/10.21956/wellcomeopenres.15094.r31563

© 2018 Shore D et al. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **David Shore**
Department of Molecular Biology, University of Geneva, Geneva, Switzerland
Maksym Shyian
Department of Molecular Biology & Institute of Genetics and Genomics in Geneva (iGE3), University of Geneva, Geneva, Switzerland

Briefly, this study presents a single negative result (absence of *orc2*, *cdc7*, of *dbf4* rescue by *rts1Δ*) already available for the scientific community in Costanzo et al., 2016. The authors should provide additional information or perform experiments to confirm reproducibility of the data, especially of the only novel finding they report.

Comments:

1. Page 3, lanes 15-16 of the Introduction: '…and the replicative polymerases (Polε, Polδ and Polα) (Yeeles et al., 2015)' is a speculation and should be corrected, as Cdc45, Sld2, Sld3, Dpb11 are not known to recruit polymerases alpha and delta. Moreover, the biochemical study Yeeles et al., 2015 used as a reference did not include polymerase delta in the reaction at all.

2. Materials and Methods: transformation of yeast is mutagenic and subsequent selection of transformants may lead to the isolation of unnecessary background mutations affecting the growth of the strain. Therefore, it is advisable to do transformations in diploid strains with subsequent isolation of haploid combinations by sporulation and tetrad dissection[1]. Thus, the authors should clarify whether the *rts1Δ* and *rif1Δ* transformations were done in haploid or diploid strains. Were haploid deletions obtained by sporulation of heterozygous diploids obtained by transformation? If so, how many different spore colonies were assessed for the ts phenotypes in each case to make sure that the growth phenotypes are reproducible and consistent? (Diploid mutant strains are indicated in the strain list, but their genotypes are not correctly described so as to indicate whether they are heterozygous or homozygous mutants. Correct nomenclature should be used.)

3. Discussion, page 9, lane 7: 'HSP60' and 'Hsp60' should be written instead of 'HPS60' and 'Hps60'

4. Curiously, the Costanzo et al., 2016 study and BioGRID do not detect/contain data on negative genetic interactions of *rts1Δ* (nor *pph21Δ/22Δ*) with *cdc45-ts*. There are slight negative interactions for *rts1Δ* with either *cdc9-1* or *dpb11-1*, but these were also reported to contain suppressor mutations that might complicate the picture (see TheCellMap web site). Could it be because the above studies used other alleles of these genes? The authors may try to compare various ts alleles of *CDC45*, *DBP11*, and *CDC9* to see if there are allele-specific interactions of *rts1Δ*. Since the authors claim in the Discussion section that this is the main novelty in their study, the use of other ts alleles, degron-based alleles or the TET-off system is advised to make this point stronger.

5. As Cdc9 participates in lagging strand DNA synthesis while Dpb11 works at the leading strand, other mutants affecting leading and lagging strand synthesis should be checked in order to clarify the matter and try and place RTS1 in one or both of these DNA synthesis pathways.

6. The proposition to overexpress *RTS1* in *cdc9-1* should be directly carried out by the authors of this study.

7. Is *SGO1* involved in the *RTS1* genetic interaction with *CDC45*/*DBP11*/*CDC9*? How will *sgo1Δ* affect the ts alleles of these genes?

References

1. Dunham MJ, Gartenberg MR, Brown GW: Methods in yeast genetics and genomics, 2015 edition: a CSHL course manual. 2015.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
No source data required

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Molecular genetics, gene regulation, chromatin, telomere biology, budding yeast

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 22 March 2018

https://doi.org/10.21956/wellcomeopenres.15094.r31562

© 2018 Fischer G. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Gilles Fischer
French National Centre for Scientific Research (CNRS), Institute of Biology Paris-Seine, Laboratory of Computational and Quantitative Biology (LCQB) UMR7238, Pierre and Marie Curie University (UPMC)-Sorbonne University, Paris, France

The hypothesis that underlies this work is that *RTS1,* the regulatory unit of the PP2A phosphatase, could counteract the initiation of DNA replication through dephosphorylating initiation factors, in a way similar to the action of *RIF1*-PP1 on MCM. However, this hypothesis is very fragile as it solely relies on a positive interaction between *rts1-delta* with a *ts-orc2* mutant, which was previously reported in a large-scale genomic screen for genetic interactions (Constanzo et al., 2010). Firstly, the score actually reported in the CellMap website (thecellmap.org) for this positive interaction is very low (0.07), below the default confidence threshold set for this study (0.08). Secondly, such large-scale genomic screens are often populated with a large number of weakly supported false positives. Therefore, it is not very surprising that a careful examination of an originally weakly supported interaction could lead to a different result. Thirdly, the authors also mention that more a recent screen reported opposite observations (enhanced phenotype in some replication mutants). Finally, the authors also indicate in the introduction that Rts1 is required for Cdc45 loading onto chromatin, which is not very compatible with their initial hypothesis that Rts1 could counteract replication initiation. Altogether these elements show that the initial hypothesis onto which the manuscript is constructed is not really reliable. I would suggest reformulating the manuscript and presenting equitably to two alternative hypotheses in the introduction, namely Rts1 is promoting or opposing replication initiation. I would also clearly state that Rif1 is used as a control to discriminate between the 2 hypotheses because as it is now presented, the reason why a parallel is made between the role of Rif1 and that of Rts1 is not really evident.

I have few other minor concerns. I found the title to be a bit misleading because this work does not directly assess the role of *RTS1* in DNA replication initiation per se but rather at the genetic interactions between this gene and other initiation factors.

I had a hard time trying to understand the logic of Fig 1. I think the legend could be more explicit to guide the reader through the different cases.

I would also suggest to replace throughout the text 'synthetic lethality' by 'negative interaction' because synthetic lethality is generally considered as an extreme case of a negative interaction where two mutations, each causing limited fitness defect on their own, result in a inviable phenotype. I would also replace 'synthetic rescue' by 'positive interaction' for the same reason. This would allow avoiding turn of phrases such as 'greater synthetic lethality'.

In the first paragraph of the result section, the authors mention that an orc2-1 mutant was reported to show growth deficiency even at permissive temperature. However, the spotting assay in Fig. 2A does not show any growth defect at restrictive temperature for the single orc2-1 mutant.

In its entirety, this study confirms a number of phenotypes previously described, including the opposing actions of *RIF1* and DDK on replication initiation, which confirms that the experiments are carefully performed. It also highlights some new observations such as a small positive interaction between *orc2-1* and *rif1.* The controls are adequate (comparison of WT and all single and double mutants) and the described data are convincing. The introduction section is really well written and give a very good overview of the interplay between the various actors involved in replication initiation. All in all, the study is a valuable additional contribution to the field of DNA replication.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
No source data required

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.